What Works:
Evidence-based Practices in Public Service Design

Ilari Laitinen

What Works:
Evidence-based Practices in Public Service Design

Ilari Laitinen

Master's Thesis

2017

What Works:
Evidence-based Practices in Public Service Design

Ilari Laitinen

May 2017

In partial fulfillment of the requirements for the Master of Arts
Degree in International Design Business Management (IDBM)
at Aalto University School of Arts, Design and Architecture.

Supervisor: Oscar Person

Abstract

This Master's Thesis examines the role and use of evidence in the context of service design in the public sector. The goal of this study was to map out and outline the views and experiences of service designers working within the public sector, about the use of evidence and evidence-based practices. Service design consultancies are working with the public sector more and more, creating value together with the decision makers and creating better and policies with decreasing financial resources. Service designers have skills, awareness, methods and know-how needed to navigate this sometimes very conservative field. The human-centric methods of service design can be disruptive for the public sector that's full of different interests, power and politics.

The study examined, how service designers make use of evidence, how evidence is defined within this context in the first place, and discussed, how evidence could be used better in public sector projects, regarding for example measuring functionality, scalability or impact. The study was conducted using qualitative research methods, and judgement sampling was used as the sampling method. This approach was chosen in order to gain information on the perceptions and representations of the use of evidence in the context of public service design. Data collection was completed during Spring 2016 using semi-structured interviews with 20 experts in order to explore and investigate how evidence is used in practice. The interviews were primarily directed to service designers working for private service design consultancies. Also experts working with the public sector were interviewed for background purposes.

In order to gain a global overview and a multicultural perspective, the informants were selected to include different countries, including Finland, Sweden, Denmark, the United Kingdom, New Zealand and the United States.

The findings describe the role of evidence from service designer's point of view, which included in practice what constitutes evidence, how and why it is utilised. The study shows that evidence has different meanings and definitions, and it can be utilised case-by-case during the process, depending on the subject matter. Argumentative evidence is the most basic form of evidence that allows service designer to show rigour in regard to design practice – and at its simplest form demonstrates the level of rigour in the service design process. When defining the problem and the solution, evidence can be empirical experience-based evidence or qualitative evidence. Promising evidence can be collected and generated through fast experiments and service prototypes, and used to validate hypotheses, or justify the relevance of the potential solution to the context. The study also showed, that the use of different indicators are metrics is useful tools for measuring the impact, and that the quality and strength of the evidence are largely dependent on the nature, scope and goals of the service design project. This study is targeted to first and foremost people working with public service design, like service designers, public servants and other public sector officials. There are recommendations at the end of this study, that can help the public sector service design process and the hands-on work. The recommendations include e.g. experimental development as well as utilising quantitative methods together with qualitative methods, to acquire stronger and more rigorous evidence base.

KEYWORDS:

Service design, public sector, evidence, evidence-based practices, experimentation, evaluation, impact, KPIs, indicators, metrics

Acknowledgement

The topic is a result of almost two years of foundational research work on the particular topic of how evidence-based approaches are currently seen as the key drivers for service delivery and policy-making in the public sector. Choosing this research topic was definitely a challenge for me. Evidence-based practices have not been studied much in the context of traditional design research or service design, and required me to deepen my knowledge and understanding of many fields that I was not previously familiar with; scientific methods, social sciences, behavioural economics, and even psychology.

My interest in this topic is a result of both personal and professional experiences: justifying decisions in an epistemically credible way and measuring the success and impact within design practice has always been seen as a difficult task. The successes and failures of my previous projects have led me to seek the definition of evidence within the context of public service design, and evaluations that show some evidence of effectiveness and at least plausible evidence of feasibility, sustainability, and transferability. As a practitioner of service design I am interested in knowing not just the '*why's*' but also the '*how's*' and '*what's*'. Asking these questions have shaped my personal learning experience and guided me through the foundational research process.

I am thankful for Oscar Person for supervising this thesis – your insights and perspective were more than valuable. I would also like to show my sincere gratitude to Seungho Lee for sharing his experiences on the topic and giving me constructive comments on my research, and professors Katja Hölttä-Otto, Peter McGrory and Ramia Mazé for giving me guidance and advice during the process. Also, accomplishing this thesis would have never happened without the contribution from my informants, so I want to thank all those who devoted their time and expertise.

A big '*thank you*' also goes to the many people who have been a part of the journey in writing this thesis: Jaakko Kalsi for always giving me something on which to reflect, Merja Lang for pushing me to finish my research, Miro Nurmela for being my scientific advisor, Pekka Perunka for giving me food for a thought and Katharina Schilli for helping me to get started and for always keeping me on track. I also want to thank my dear colleagues and friends from the IDBM and ME310 programmes for being my peer support network. I also want to thank all my fellow colleagues from the Design for Government (DfG) course – this thesis would have ended up looking very different without your support, experience and knowledge.

Lastly, I would like to thank my family members for their encouragement. And special thanks to my beloved one Enni, who has kept me grounded through the entire process, and ensured that this book reached the stage of completion.

Table of Contents

1

BACKGROUND

Public services have been shaped by countless political, social, economic and historical factors. In addition, different governmental, institutional and structural reforms have influenced and changed the nature of public sector management. Most recently, the technological revolution driven by the emergence of smartphones has dramatically changed the way public services are operated, produced and delivered. The case study of Estonia indicates how rapidly a small nation with an active, digitally-aware leadership can advance. Now, numerous Estonians consider their online services more secure and advantageous than traditional strategies of dealing with the government.[1]

1 Scott 2014

The public sector is one of the largest sectors in most economies and faces a major challenge in providing the kind of services that citizens today expect and demand, with increasing financial resources. This is to say nothing about meeting the major challenges facing us in the future, such as the environment, increasingly marginalised groups, and the ageing populations.[2] This is where public sector innovation typically falls short, and service design comes into play – aiding the public sector in becoming more agile, productive, comfortable with risk-taking and responsive to user needs.

2 Snook & DMA 2014

There is a general development today over the public sector towards more noteworthy responsibility, transparency and viability. To enhance these issues, various actors, for example, government, public sector organisations and private sector actors are striving to bridge the gap between action and knowledge.[3] There has been a tendency towards evidence-based policy-making, that is extensively proposed to better direct resources and procedures towards results.

3 Miller & Rudnick 2012

In 1999, The British government introduced another vision for the role of government in the United Kingdom, calling to attention that it needs to improve the utilisation of evidence and research in policy-making and better concentrate on policies that will convey long-term objectives. Also, the government emphasised the need to learn from experience:

> *'Government should regard policy making as a continuous learning process, not as a series of one-off initiatives. We will improve our use of evidence and research so that we understand better the problems we are trying to address. We must make more use of pilot schemes to encourage innovations and test whether they work. We will ensure that all policies and programmes are clearly specified and evaluated, and the lessons of success and failure are communicated and acted upon. Feedback from those who implement and deliver policies and services is essential too. We need to apply the disciplines of project management to the policy process.'* [4]

4 Cabinet Office 1999a

Politicians, public servants, government bodies, and official entities around the world have now followed the UK's lead in embracing design as an means to shape public policies and services – and are learning to acknowledge the value of providing their populations with what they need. Citizens today understand the value and importance of their own voice and beliefs, and do not shy away from having their needs met. It is this same force, along with the need to see immediately effective results, that drives service design to produce positive results in the public sector. The public sector needs to keep pace with the speed of globalised world and rising public expectations, if it is to serve us well into the future – as well as the availability and potential of new technology to deliver better services at a lower cost.

1.1
Framing of the Topic

This Master's thesis aims to answer questions about how the concept of evidence relates to public service design, and explores the understandings, meanings and the use of evidence from the service designer's point of view. The purpose of the study is to map out the role of evidence and evidence-based practices in the context of public service design.

Evidence-based practices are used to produce *evidence* – knowledge that is utilised to support a hypothesis, perception, claim, or decision –, and with regards to the public sector, to support decision-making, scale-up, implementation and replication. This study seeks to outline a range of methodologies for identifying and applying these practices into complex and ambiguous challenges that the public sector faces.

The study aims to characterise the process by which findings, information, or data becomes evidence. The study also tackles service design practitioners' challenges on evaluating their hypotheses and justifying the findings to themselves and others, in this case public sector organisations. Also, the study examines and explores different strategies for measuring service outcomes using field experiments and interventions. Typically, these experiments last for a defined period of time, and outcomes can be measured in terms of changes in desired outcome indicators before and after the experiments.

While different service design methods generate evidence, this data is synthesised and managed in quite different ways compared to evidence collected in the so-called hard sciences.[5] There, evidence collection relies on quantitative, statistical data and trials, whilst service design methods place an emphasis on qualitative data, which is rich in context and focused on the particular experience of individuals.[6] By understanding how service designers experiment, measure, evaluate and justify, best practices may be found to apply these principles in practice.

5 Carr et al 2011

6 Hagen et al 2014

1.2
Key Research Gaps

To give valuable services to the public, it is imperative to concentrate on how these services are designed and implemented.[7] Whether the issue is introducing economically viable medical facilities or reconstructing public transport, service design fully measures public opinion and needs before developing a solution.[8] Mager (2016) emphasises that the service designers should deepen and broaden their knowledge about the public sector in order to face the specific opportunities and challenges in this domain.[9]

While research on service design has increased in the past few years, key gaps in knowledge remain. Mager sees the need for expanding and developing the in-depth knowledge about service design in the public sector.[10] There is a wide array of literature about service design, but the great majority of the current research is mainly focused on different tools and methodologies, as well as earlier stages such as involvement strategies, idea generation and prototyping.[11] While there is a growing amount of literature on using evidence, especially in social policy and policy making, literature on the concept of evidence-based practices in the context of service design is still almost non-existent.[12]

There is amazingly little research on the way toward using evidence within design practice which describes how the qualitative and experimental research methods stand in an epistemically credible way to the information and data

7 Tuulaniemi 2011
Meidutē et al 2011

8 Mager 2016b
IDEO et al 2016

9 Mager 2016a

10 Mager 2016a

11 Blomkvist 2011

12 Breckon &
Dodson 2016

they deliver. Applying experimental and evidence-based methods to policy making has been introduced on a theoretical level in the 2015 report for Prime Minister's Office in Finland.[13] Most recently, Codinhoto (2013) and Sustar and Feast (2016) have addressed the requirement for better understanding of use and nature of evidence in design practice.[14] As service design grows as a discipline and design methods become more widespread, there is a great need to strengthen its validity through evaluative frameworks that provide a stronger evidence base of its role in innovation.[15] The requirement for more evaluative frameworks is also to the issue of service design's legitimacy, professionalization and codes of practice.[16]

The collection of evidence is an integral part of the iterative service design process, which is based on the '*test, learn, adapt*' philosophy.[17] Ideas need to be tested and then validated in order to continue the process – and, this way evidence is generated along the way. There is lack of studies on the '*evidential tapestry*' service designers can provide for when justifying design decisions to clients and stakeholders, or when showing the potential impacts of the service. The collection of evidence is essential for the evaluation process, and measures of impact and effectiveness are essential for understanding the outcome of a service.[18] Evidence also relates to delivery quality and fidelity to the service model. In order to evaluate the overall success of the service, robust evaluation methods should be applied at various points through the service design process.

13 Demos Helsinki et al 2015

14 Codinhoto 2013; Sustar & Feast 2016

15 Sangriorgi et al 2014

16 Kirchberger & Tether 2014; Collins & Cook 2014

17 Cabinet Office & BIT 2014

18 Salmelin 2014

1.3
Research Questions

The main research question is:

- **What are the evidence-based practices in public service design?**

The following sub-questions elaborate on it:

- **What constitutes evidence in public service design?**
- **How is evidence synthesised and used to inform decision-making?**
- **What is the role of evidence in the different phases of the design process?**

In order to fully answer the main research question and the sub-questions, the following question will form a starting point for my empirical work:

- **What is the definition of 'evidence'?**

1.4
Research Framework

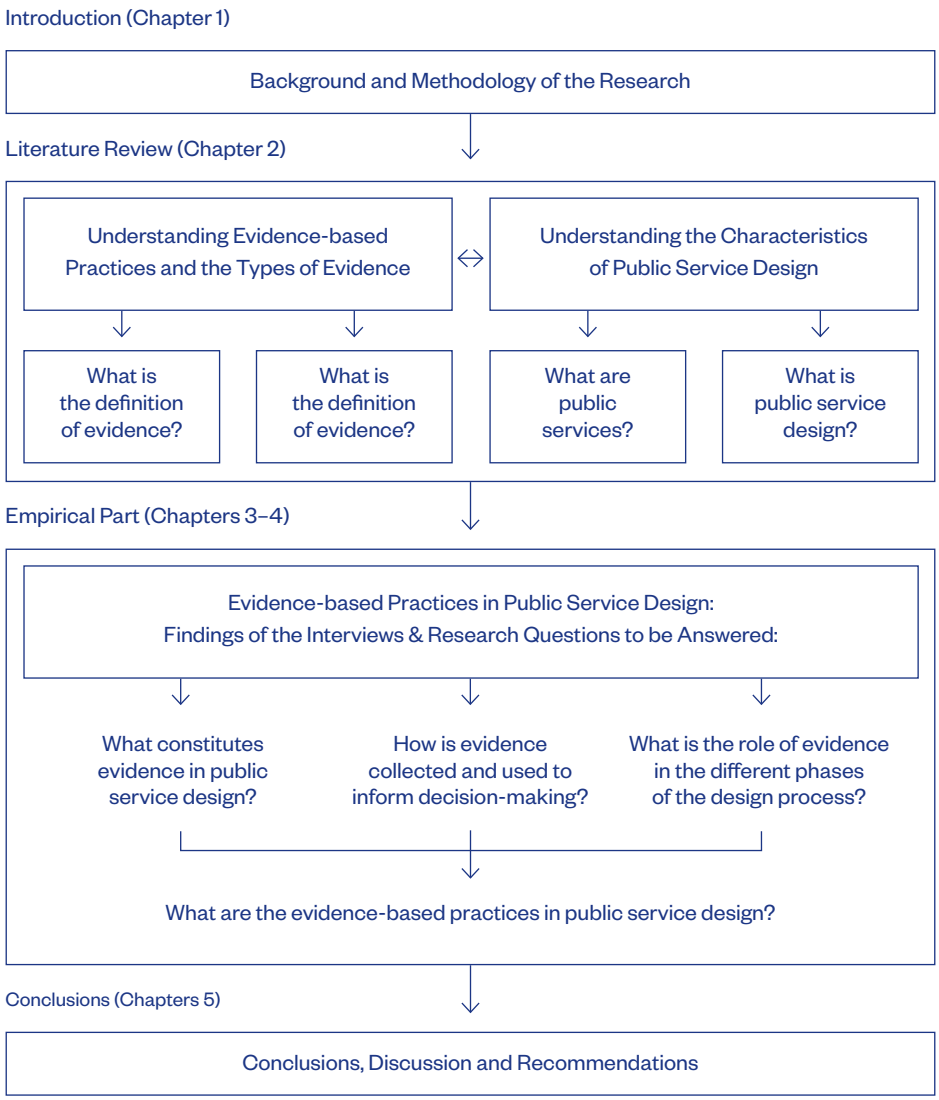

FIGURE 1 A schematic research framework and thesis outline.

1.5
Research Methodology

The research outlined in this Master's thesis aims to answer the proposed research questions and sub-questions. The research is divided into three parts: 1) *literature review*; 2) *empirical part* that consists of the findings of semi-structured interviews with service designers and other experts, and 3) *thematic analysis* that examines and outlines the different evidence-based practices in public service design. A schematic research framework is visually represented in **FIGURE 1**, describing the outline of this thesis.

Literature review is the first part of the research with the secondary data derived from relevant books, journals, articles, conference proceedings, and reports. The purpose of this part is to figure out what kind of methods are important in gathering and evaluating evidence, what are the different conceptual frameworks of evidence-based practices, and the different empirical foundations for this research project. The currently available literature on the use of evidence-based practices across disciplines is investigated, drawing on the principles of empathic design, behavioural economics and social sciences. Although literature on evidence-based practices is very rich, only few works deal with the design of public policy or public services. There is practically no literature dealing specifically with this topic. Data on the evaluation and impact of service design is also discussed, but unfortunately the amount of sources is limited as well.

The second phase involves the collection of primary data, where the information is collected through semi-structured expert interviews.

The empirical part of this thesis consists of semi-structured interviews with service designers and those in the design industry with years of experience working with the public sector. Also experts in the fields of public policy, behavioural and social sciences were interviewed for background purposes – to get specific and detailed knowledge regarding existing practices and experiences working with the public sector services in Finland and also other countries. The purpose of this part is to capture and explore the use of evidence in practice, reveal the benefits and disadvantages of the different evidence-based practices as well as examine how these practices work in different organisations across the public sector in different countries. Interviews were an appropriate method of gathering data within practical approach.

In the last phase, a thematic analysis is carried out – outlining the use of evidence-based practices in the public sector service design through the analysis that reflects the similar and different ways of working, and illustrates previously unrecognizable issues concerning evidence-based practices. Analysis of data received from interviews helps to describe the role of evidence from service designer's point of view, and whether they have adopted evidence-based practices in their work – including what constitutes evidence in the first place, and how and why evidence is used.

It is important to emphasize that based on the nature of this study, time factors and other limitations, the thematic analysis is primarily built on the semi-structured interviews.

2
LITERATURE REVIEW

This chapter provides a theoretical background on the terms and concepts that are relevant for this study. As the title suggests this study is concerned with *evidence* and *evidence-based practices* in the context of *public services* and *service design in the public sector*.

The characteristics of these key terms and concepts are defined in the following pages, as they are used throughout the thesis.

2.1
Public Services

The European Parliament has defined public services as '*services of public interest or public utility, such as electricity, gas and water supply, transport, postal services and telecommunications*', and '*economic activities of general interest set up by the public authorities and operated by them or by delegated separate operators (public or private)*'.[19] Public services are services provided by the government to citizens that live within its jurisdiction, making the state visible to its citizens through the public sector. Public services are built around the citizens' essential needs, and the public sector is expected to react to these needs with requests for better roads, social arrangement, policing and other public services. There is a social consensus that states are universally expected to provide certain services to its citizens, regardless of the social status, physical or mental abilities.[20] One clearly stated value of public services is, that they are extremely substantial, concerning its acute value in the citizens' everyday lives.[21] According to Rotberg (2014), public services are a visible connection between what citizens give to the state (*taxation*), furthermore what they expect consequently (*some level of prosperity*). Public services also offer the '*social contract*' between citizens and public sector, therefore being an essential dimension of state-society relations.[22]

19 Leinonkoski 2013

20 Whaites 2008

21 McLoughlin 2015

22 Rotberg 2014

In general, public services are provided by the government on a non-profit basis. The government has traditionally been responsible for setting and formulating the behavioural rules – by regulating laws and policies – while

simultaneously being a major player in the system – by providing services for the society.[23] Although semi-public entities, public-private partnerships and exceptions do exist, public sector services are not driven by profits or commercial goals, but rather by missions focused on serving citizens.[24] These can cover everything from social services and healthcare to environmental protection and education. According to Murray (1990), the public sector has some specific characteristics that set it apart from private and commercial entities, and there is less of the assertion that public affairs and private business should be managed the same way. As Murray has expressed it, public servants need to be careful while rebutting the claim not to give the impression that efficiency is either compromised or unsatisfactory in the public sector – individuals who recommend private sector efficiency as the cure for public sector inefficiencies are signifying that they are not much aware of the constraints and objectives of the public sector.[25]

23 Carmichael & Kaufmann 2001

24 Gross & Smith 2015

25 Murray 1990; Humphreys 1998

Public services can contrast altogether from commercial and private sector services in different ways, as outlined in **TABLE 1.**[26] When the capability of private sector services is evaluated to match people's needs, we are looking at things like buying behaviour and consumption patterns. In the case of public sector services we are talking about the public interest, '*the common good*', or the long-term impacts of the service – mostly because the purpose and nature of budgeting for the public sector is fundamentally different, as opposed to the private sector.[27] Also, what makes a huge difference here is the fact that, sometimes, customers do not have a choice of alternatives to these highly regulated public services.[28] Public services can have wide social and economic impacts. For example, public services that focus on people's health and general wellbeing such as social services and education services, can have a real impact on disease progression and hospital admissions. In spite of the fact that the figures for success and failure over the public and private sectors might be practically identical, there are some contention that the public sector should aim for higher level of progress than the private sector because the government is publicly funded.[29]

26 Pearce 2000

27 Leinonkoski 2013

28 Leinonkoski 2013

29 Pearce 2000

PRIVATE SECTOR	PUBLIC SECTOR
Measurable financial objectives.	Have multiple aims, so hard to measure success. However, there is some argument that the public sector should aim for higher rates of success than the private sector because government is publicly funded.
Business driven by competition.	Limited competition – generally not in competition with the private sector.
Business driven by competitive advantage gained by the quality of service.	Various objectives, well-being and economic growth among others.
Often not visible to the public or shareholders.	Highly visible to the public and the media.
Open to risk taking.	Managed in a risk averse culture.
Designed to limit damage when they are in difficulty.	Difficult to adapt to change because of scale and complexity.
Less constrained by legislation and regulations.	High complexity – constrained by legislation and regulations.
Financed by charging users.	Financed by tax revenues – users range from listed companies to grannies with dementia, and even future generations.
Increased acquisition of new customers, customer loyalty and retention of customers are essential.	The service provider may not want you to need the service at all (or at least not twice) – such as drug-addict rehabilitation.
Tendering process can be used in order to maximise competition, providing an opportunity for a large pool of suppliers to make an offer, as well as having a greater choice in selecting a supplier that offers value for money.	Many countries are legally obliged to release tenders for works and services. Typically the goal of the process is to obtain the cheapest possible price for a project with fixed and predetermined scope and quality. The tendering process can be a long and complex process with many constraints.

TABLE 1 A comparison outlining the differences between private sector and public sector services. Adapted from Pearce (2000).

2.2
Service Design in the Public Sector

The aim of this chapter is to give an overview on, why governments need to innovate and how service design can be an effective approach to public sector innovation helping public institutions to transform the process of policy-making and public service design.

Service design is a practical, design-based approach to deal with innovation, giving a systematic way to public sector innovation. Service design can be portrayed as a toolbox, a set of techniques, tools and methodologies that can be made use of, at various phases of the innovation process to support the benefit of existing services, additionally address societal issues at a systemic level. Service design has an ever-growing role in the public sector, taking a human-centred approach to implement and formulate policies, making an incentive for society through the formation of new innovative services.[30] Service design connects the users outside of an organisation – citizens, customers and the wider population – with those enabling the service delivery – public servants, policies, managers, internal users, processes and systems.[31]

30 Bason 2010b

31 Snook and DMA 2014

While the public sector is considered less amenable to innovation, there has been a great deal of interest in recent years in the application of design-driven innovation in the public sector, and creating better policies with decreasing financial resources.[32] Even the European Commission has

32 Borins 2001

recognised design as a key driver of service innovation and human-centred innovation. However, these human-centred aspects of innovation are still deficiently coordinated into innovation policies.[33] Economic sustainability has been cited as a primary variable that public servants must unerringly consider, time and again, when both crafting and implementing a project. All public sector initiatives are intended to significantly improve the quality of life for the public. Thus, target groups must be identified and, consequently, prioritized. Therefore, service design respects a delicate balance between stakeholders, interest groups, and affected demographics that needs to be maintained.[34]

33 Commission of the European Communities 2009; 2013

34 The Interaction Design Foundation

In 2016, Service Design Network, the leading professional organisation for service design with approximately 30,000 affiliated members and followers, published a comprehensive report which examines the role of service design in the public sector.[35] Published in October 2016, '*Service Design Impact Report*' provides an overview of how service design is currently contributing to the innovation of public services. The report is based on a global survey with service designers and public servants as well as interviews gathered from 17 experts in the field.[36] The study focuses on identifying and analysing five key areas where service design is contributing in the public sector: *policy-making; cultural and organisational change; training and capacity building; citizen engagement;* and *digitisation*. These themes are further explored in the following chapters.

35 Service Design

36 Mager 2016b; Service Design

2.2.1
Policy-making

In the '*Design for Public Good*' report published by SEE Platform and UK Design Council, the authors propose the three-step Public Sector Design Ladder (**FIGURE 2**) to illustrate how different public sector bodies utilize design, and secondly it serves also as a roadmap for progress indicating where design-led methods can be utilised.[37] The first step – *design for discrete problems* – indicates that design used to tackle discrete problems, meaning that the projects are one-of-a-kind thus prevents the development of commissioning organisation's design capabilities. The second step – *design as capability* – indicates that the public sector organisation understands the value of design-led methods and are able to solve problems without hiring an external service designer or consultancy. Finally, the third step – *design for policy* – indicates that design is utilised as an approach to innovation in policy-making.[38]

37 SEE Platform 2013

38 SEE Platform 2013

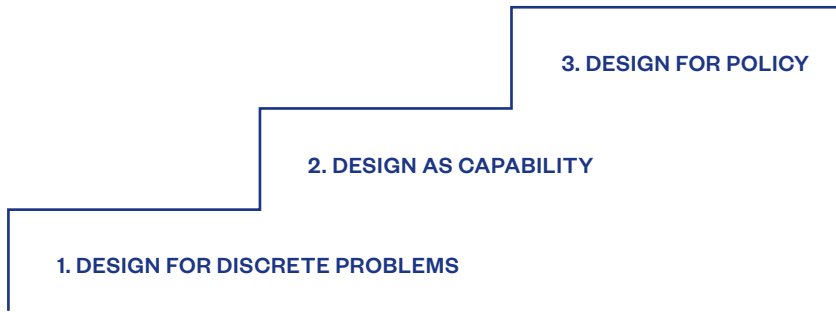

FIGURE 2 The three-step Public Sector Design Ladder illustrates how different public sector bodies can utilise design. Adapted from SEE Platform (2013).

Both industrialised and emerging economies have been moving to lead the way in adopting new policy tools to reinvent their public services. Design-led innovation have become an increasingly popular phenomenon in policy making, gaining acceptance far and wide. Over the past decade, a number of countries have started to integrate design into public policies and services as part of their national strategies. Various governmental and non-governmental initiatives, innovation units or '*labs*' inside public institutions, such as the *Helsinki Design Lab* and *Design-Driven City* in Finland, *MindLab* in Denmark, *Behavioural Insights Team* and *Policy Lab* under the UK Cabinet, *Laboratorio de Gobierno* in Chile and the *Public Policy Lab* in the United States, have been set up to bring an experimental approach to policy-making, addressing systemic challenges facing governments and citizens.[39]

39 Bason & Schneider 2015; Design for Government 2017; Laboratorio de Gobierno 2017

The work that is being done at these initiatives show that systemic change can occur for the benefit of the public with the assistance of design-led innovation.[40] In his article, '*The Value of Service Design for Public Policy Making*', Nicolas Rebolledo, a lecturer on service design at the Royal College of Art in the UK, seeks to address the question of why government should innovate. In arguing for a review, the government, he noted three reasons for innovation: [41]

40 Mulgan & Leadbeater 2013

41 Rebolledo 2016

1. To *do more with less*; increase the quality of services to meet with the increasing demand and expectations of citizens in time of scarce resources, summarily, it is about improving productivity and effectiveness in the process of policy and service design and delivery. This imperative calls for a more integrated approach for the development of policies and services.

2. It is necessitated *in a bid to regain trust in public institutions*; making government more open. This would aim to resolve legitimacy crises, cited the reduction in trust for the government after the crises in Europe. This view was also echoed in a report by ERCAS.[42] All with the objective of encouraging a more participatory process and operations that strengthen relationship between government and the people.

42 ERCAS 2015

3. It is needed in *dealing with the complexity of public problems*, there is an urgent need to rethink the traditional policy approach and embrace the nascent systems views. Lending a voice to this argument was Muir and Parker (2014) pointing to the rigid nature of traditional policy approaches used in government.[43]

43 Muir & Parker 2014

All of the aforementioned drivers, forms the construct for a multi-dimensional innovation imperative for the public sector. Lastly, Rebolledo appraises the value of service design in policy formulation arguing that the change of paradigm to the reinvention of policy design would contribute to service design through, change of orientation, more practical approach for policy design and providing an alternative language for communication.[44]

44 Rebolledo 2016

According to Vaajakallio et al (2013), creating and applying a human-centered design approach in the public sector can be considered a radical innovation, requiring a major adjustment in culture and conduct, time, assets, and management support.[45] Bason (2015) has emphasised that moving towards public sector innovation suggests particular difficulties and tasks for public leaders at all levels, lawmakers, executives, mid-level directors and the heads of organisations.[46] Service design serves an invaluable purpose when applied to policy-making, improving the efficiency of public services and establishing effective communication between the general populace and the official bodies in charge of implementing change. Service design often encounters innumerable challenges and obstacles while surveying the public in order to construct a model that can be socially, politically, and economically viable for each individual.

45 Vaajakallio et al 2013

46 Bason 2015b

2.2.2
Cultural and Organisational Change

In his article, '*Embedding Design: Towards Cultural Change in Government*', Jesper Christiansen (2016) is commenced by building a case for the government, arguing that it is an institution which by legal provisions has the power to control humans and organisation in a given territory.[47] In addition to this is the jurisdiction to making and enforcing of laws capable of aiding the performance of its responsibilities and employing instruments of coercion when and where necessary. Hence, man and his environment are indispensable to any government's existence. These two key elements also affect each other in various ways; as the environment influences man, man also exerts a great deal of influence on the environment. Given the complexities of organisational structure, culture change is a significant challenge.[48] It is therefore not only rational but also imperative for government to be flexible in its approach or method it employs in solving the unique problems attached to each stage of human and environmental metamorphosis.

47 Christiansen 2016

48 Cummings & Worley 1997

This problem solving approach of the government and its institutions is what Christiansen addresses, bemoaning the rigid approach of government towards solving these problems. Design is a pivotal tool in engineering the desired changes in social environments and policy making context.[49] Also described as a core, indispensable and enabling factor to the approaches government uses in driving cultural change in the environment, Christiansen mentions *goal concentration, citizen engagement, system consciousness* and *demonstration of ideas on a smaller scale* before full implementation as the key attributes of a design system. Public organisations cannot only adopt a system of survival by adjusting to their condition, as the reason for a public servant is actually to shape the environment.[50] Christiansen argues that there is a need for government to recognize and employ design in its methods and approaches both vertically, between the state and local level, and horizontally, across

49 Christiansen 2016

50 Mulgan 2009

government ministries and institutions. Examples of different campaigns and countries embracing this strategy include, but not limited to:

- *MindLab*, an internal cross-governmental design lab in Denmark, utilizing design techniques for finding better approaches for addressing issues and designing the fitting procedures to form new thoughts into practical outcomes.

- *e-Estonia*, the term generally used to portray Estonia's emergence as a standout amongst the most exceptional '*e-societies*' in the world; the association between a tech-savvy population, a proactive ICT sector and a forward-thinking government.[51]

- *InWithForward*, a social enterprise based in Canada, specifically working with individuals and organisations to change results and behaviour through peer-to-peer influences, experimental co-design and social movements.[52]

51 Republic of Estonia 2014

52 Christiansen & Schulman 2015

Consequently, design system should be a standard for government in determining its approaches in order to achieve set goals and objectives.

Katz and Gott (2016) argue on the rise of demand of design-led approaches to organisational change, and also make a case for service design in addressing organisational challenges. Based on the reasoning that service design favours a wider approach to problem solving that is collaborative, human-centred, data-driven, emergent and iterative, sensorial and multi-dimensional. Yet it focuses on how the wider service design approach is applied to organisational and systemic challenges that are complex and nonlinear.[53] Further, it puts forward a taxonomy of challenges faced by public sectors into wicked or complex, and makes a constructive case by arguing that the problems in the public sector are system challenges, which sharply contrasts with linear and more familiar changes.

53 Katz & Gott 2016

According to Katz and Gott, the ways these challenges are too often met derive from the metaphor of a machine, *re-design, re-engineer, leverage* – a metaphor that has its origins in the early 20th Century Taylorist views of effecting change.[54] Narrated how weak Taylorists approach is less efficient in making the desired change because the better part of it is based on five shaky premises. As argued by Fuda (2009), these premises are; people are objective, change can be accomplished, we have the well-known take-off point for change, there are X steps to change and that change, itself, is the goal.[55]

55 Fuda 2009

In order to address the multitude of citizen and private sector expectations, the public sector needs to become more open, innovative, collaborate, agile and participatory. This, however, can be a difficult and slow process. Adopting a complexity and living systems perspective demands a very different mindset and way of working and importantly, service design provides a philosophy and discipline to address complexity and change in living systems. When seeking to address issues such as the growth of long-term conditions, homelessness or poverty, Katz and Gott suggest that two key perspectives must be adopted first of which is complexity, and importantly, *the differentiation of complicated and complex problems*.[56] Both sharply contrasts in that, in the case of the former right answers are undiscoverable as against the latter, where at least one right answer exists – and should be greatly considered when intervening. The second perspective alludes to the fact that the spaces into which we step on behalf of clients are living systems, and far removed from the machine metaphor and the false certainty and optimistic predictability implied. This is also exemplified in the works of Wheatley et al (2012), providing further advice on how to facilitate successful organisational change.[57]

56 Katz & Gott 2016

57 Wheatley et al 2012

2.2.3
Training and Capacity Building

In their article '*From Capability Training to Capacity Building*', Yee and Choukeir (2016) re-affirm the role of service design in driving innovation and change in the public sector evidenced in a number of factors including, and *the rise of innovation labs in the governments around the world, the growing number of public sector service designs projects*, in addition to the *increasing number of academic research investigating the value of service design in the service sector*.[58] Different scholars have addressed the obstacles to service design implementation namely on the grounds of capability and capacity. Yee and Choukeir highlight a range of approaches to service design training and capability building in the public sector using reasoning from different parts of the world, expressing this same view is Brooks (2015).[59]

58 Yee & Choukeir 2016

59 Yee & Choukeir 2016; Brooks 2015

According to Yee and Choukeir, service design is seen as a cost-effective approach that maximises policy effectiveness in revamping public service, it then goes on to make a case for the importance of training and building capacity in service design. First being that, changes do not occur without building the capability and thus increasing an organisation's capacity to change. Secondly, it is indispensable that public sector involved in designing and delivering new services in the public sector are offered relevant training and support to ensure they have the capability to implement the newly designed service. And lastly, training is imperative to keep and have people with the right skills and expertise in order to ensure the sustainability of the newly introduced service.[60] Following the need for above mentioned training and capacity building, Potter and Brough (2004) outline different sources of training and capacity building in the public sector.[61] These are in no particular order, via *embedded design units*, through *intermediary organisations* that work alongside the public sector, and lastly, *external consultancies*, as described in **FIGURE 3.**

60 Yee & Choukeir 2016

61 Potter & Brough 2004

- **EMBEDDED DESIGN UNITS** are a strategic-level units in charge of specific design and governmental programmes, as well as creating organisational design capacity in the public sector. These units typically work from within the public sector. The municipality of Helsinki in Finland serves as a good example of internal innovation in the public sector's local governance: the value of design is considered high, service designers have recently been employed by the municipality, and Helsinki is also among the first cities in the world to hire Chief Design Officer.[62]

62 Mager 2016b; Helsinki Design Weekly 2016

- **INTERMEDIARY ORGANISATIONS** are cross-governmental design units that work with governmental units and other parts of the public sector on a project-by-project basis. One good example is already mentioned MindLab in Denmark, set up by the Ministry of Taxation and the Ministry of Employment, the Ministry of Economic and Business Affairs to bring together private enterprises, the research community and government under one roof to advance design-led innovation.[63]

63 Mager 2016b; Social Innovator 2017

- **EXTERNAL CONSULTANCIES** are privately owned agencies and design consultancies that provide consultancy from an independent design practice on a project-by-project basis.[64]

64 Mager 2016b

It is noteworthy to mention that experiential training where organisations apply new approaches, tools and methods to their life challenges have proven to be the most effective. However, these new ways of working can be a challenge, bringing many forms of resistance to change. Ramirez (2008) states that service organisations are considered as complex social systems that are made up of '*value systems*' – a network of stakeholders.[65] As the complexity of challenges is growing, design needs to be able to collaborate with a wider number of stakeholders to enable systemic change.[66] Vuontisjärvi (2015) suggests that it is critical to distinguish and include these key stakeholders already in the very first phases of the development process.[67]

65 Ramirez & Mannervick 2008

66 Sangriori & Pacenti 2008

67 Vuontisjärvi 2005

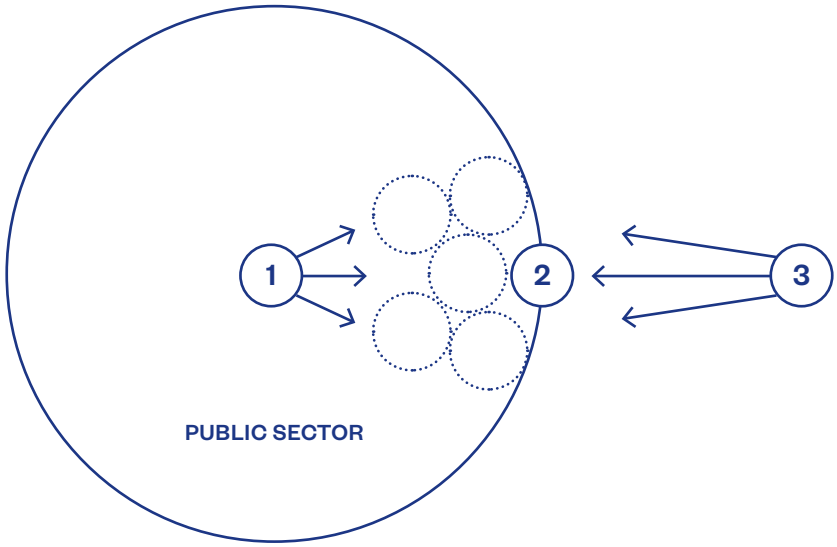

FIGURE 3 The diagram illustrates different service design entities across the public sector innovation; 1) *embedded design units*; 2) *intermediary organisations*; and 3) *external consultancies*. Based on Mager (2016b) and Design Commission (2013).

2.2.4
Citizen Engagement

Christian Bason, the CEO of the Danish Design Centre has stated that '*citizens are experts in their own lives and nobody – nobody – else can claim that role*'.[68] Involving citizens, people, communities and stakeholders meaningfully is essential to successful public service innovation. Citizen involvement proves to be crucial when planning and implementing changes to be introduced into the public sector. Bason (2010) explains that public sector innovation '*requires the courage to co-create new solutions with people, not for them*'.[69] Without this fundamental input, service designers and public servants will not be able to fully capture the nuances of what the public demands.[70]

68 Bason 2010b

69 Bason 2010a

70 Mager 2016a

Drew (2016) discusses the aim of promoting mass engagement in policy formulation through service design, and argues that service user focused on user engagement is the connexion between policymakers and service designers. Substantiating her arguments with experience in the United Kingdom exemplified by the establishment of Policy Lab to supports department to engage citizens in persecuting policy challenges. Joining a growing number of innovation teams that connects policymakers and service designers in an attempt to deliver human-centric approach to policy formulation.[71] Furthermore, Drew emphasises the importance of engaging citizens rationalised on an individual and collective basis, and outlines a range of citizen engagement methods, from *informing* and *consulting*, to *collaborating* and *empowering*. Drew also argues that citizens need to be seen as co-creators of services rather than recipient and the fact that they need to feel engaged in decisions that affect their lives.[72]

71 Drew 2016

72 Drew 2016

One example of the impact of public dialogue on policy making is the constitutional convention held in Ireland post 2011 coalition government, where 100 citizens met over a period of 12 months and discussed various challenges and general issues facing the society.

The convention achieved remarkable results thanks to public involvement, such as paving the way to the successful referendum that legalised the same-sex marriage in Ireland. Donetto et al (2015) argue that, as a necessary condition to bring in service design methods to the public sector's citizen engagement, policy makers and service designers need to work on two accounts; policymakers and frontline service providers need to be open to incorporating it into their everyday practice. Secondly, and to push the frontiers for next practice.[73]

73 Donetto et al 2015

Another argument worthy is, service designers can bring the individual citizens' experiences into these debates, by using film ethnography and documentaries of people's lives or asking participants to use different cultural probes, thus becoming user researchers themselves. The visual material of individual examples known as '*thick data*' can result in wider patterns of evidence through the power of storytelling. Not forgetting its argument for data science and digital platforms as a tool in engaging with more people.[74] Interestingly, it points to the ethical question that plagues data science relating to whether the public expect the opinions they post on social media to be considered comments on government policy. Traditionally, policy tends to be hierarchical and soiled as the public sectored is marred with bureaucratic process slowing down implementation and needs to be urgently addressed and lastly, in recognition of citizen's need to engage, policymakers also have to build the '*supply*' of citizens ready and willing to engage.[75]

74 Gage 2002

75 Drew 2016

2.2.5 Digitisation

Many governments have come gradually to understand that they need to understand the capabilities and opportunities of digital era to transform the public services. In her article '*The Future of Public Services*', Louise Downe (2016) strikes a cautiously optimistic tone when it argued that digitising a service would not necessarily remove the flaws in such service without addressing the foundational problems inherent in the service from its root. Further, Downe states that for services to be effective, every service have to begin its design with the end user in mind, asking and providing answers to the question of what the users' needs are, from the very start and designing such services with answers obtained therein. Noting that it is not just about tweaking what we do, it is more about revamping what we do and how we do it.[76]

60 Downe 2016

Modern, tech-savvy citizens who make use of social media platforms and mobile devices live their day to day lives through the lens of instant gratification. If they require a quick ride to a store, they can order one with a few quick taps on their mobile devices. Whether to shop online instantly, hire contractors for work, obtain personally-tailored financial advice, or browse housing options, nearly every single task can be achieved within a few minutes and in a handful of keystrokes. The public sector thus experience a fresh challenge: with an audience so highly demanding and used to getting immediate results, how can public servants and civic leaders best cater to the ever-evolving needs of the population in a short time-frame? To deliver effectively service, the government needs to have a shift in thinking from their traditional departmental silos ideology, as users care less about which department does what or that, they just want to get things done as seen in Fishenden and Thompson (2012).[77]

77 Downe 2016; Fishenden & Thompson 2012

The first wave of digitisation as taken us only so far, for the second phase we need to think beyond service shift and rethink government services and the government that delivers them, similar to Dunleavy et al (2006).[78]

78 Dunleavy et al 2016

Digitalisation is one of the most important means for the public sector to innovate their services. Downe (2016) argues that we are in the first phase of digitisation, evidenced by the increase digitisation of government services, justifying its reasoning with examples of such programmes which recorded tremendous success from different countries, such as the *Government Digital Service* (GDS) in the United Kingdom, *United States Digital Service* (USDS) and *The Digital Transformation Office* (DTO) in Australia.[79] In 2014, the United Kingdom hosted the first summit of the global *D5*, or *The Digital 5*, which united five of the most digitally advanced nations in the world; the United Kingdom, Israel, Estonia, South Korea and New Zealand, with the aim of strengthening the digital economy.[80]

79 Downe 2016

80 Williams-Grut 2014

In this context, it is especially noteworthy to mention, that the UK government's redesign of its digital services has been a particularly large, complex and impressive one. Tom Loosemore (2015), the former director of the Government Digital Service in the United Kingdom, describes that as a result of moving to the new '*platform-as-a-service*'-model, bringing together 350 different governmental websites onto one website, the UK government has saved billions of pounds of public money.[81] According to Whicher (2015), the project has saved £55-70m (€74-94m) just by replacing the two main government support websites, and the estimated annual savings from the shift to digital services is £1.7bn (€2.3bn).[82]

81 Loosemore 2015

82 Whicher 2015

2.3
Evidence

For a start, the goals of this thesis are directly linked to the interpretation of the term '*evidence*'. The service design process commonly begins with a target for action, where the fundamental question is, '*what do we need to know, given what we want to do?*', which guides us towards distinguishing those types of information, data and evidence that is required to accomplish the specific objectives in question.[83] As stated by Miller and Safer (1993), '*evidence*' is an vague term, and there have been endeavors, without finish accomplishment, to sought a meaning of the term that will fulfill the necessities of each circumstance in which the term is utilised.[84] The term evidence is associated with the words '*information*', '*knowledge*', '*observation*' and '*data*', sometimes even '*proof*' or '*fact*'.[85] The distinctions and relationships among there terms need to be carefully examined when they are related to the concept of evidence. 'Information' is not the same thing as '*evidence*'. The two terms are connected in imperative ways, however they are not interchangeable.[86]

83 Miller & Rudnick 2012

84 Miller & Safer 1993

85 Merriam-Webster 2017; Oxford English Dictionary 1971

86 Miller & Rudnick 2012

We can consider data just as the realities or subtle elements found out about something or knowledge produced through study or experience. It can also be tacit, in people's heads, or explicit, in documents – electronic or manual. Evidence, however, is that information which supports, justifies or contradicts beliefs, hypotheses, theories, or claims.[87] According to Achinstein (2011), scholars have developed different theories and definitions of evidence that have at least two purposes: 1) to clarify what it means to say that some fact is evidence that a hypothesis is true; and 2) to help determine whether (and to what extent) reputed evidence supports a hypothesis.[88]

87 Goodman & Royall (1988); Sackett et al 1996

88 Achinstein 2001

While evidence is constantly comprised of information, information all alone is not evidence. Information only becomes evidence when it is connected in an evaluative and the logical procedure by the designer or researcher to affirm, verify or invalidate a claim. **FIGURE 4** describes one way how information yielded from research is used as evidence, that is further used to build situated theories to formulate design propositions.

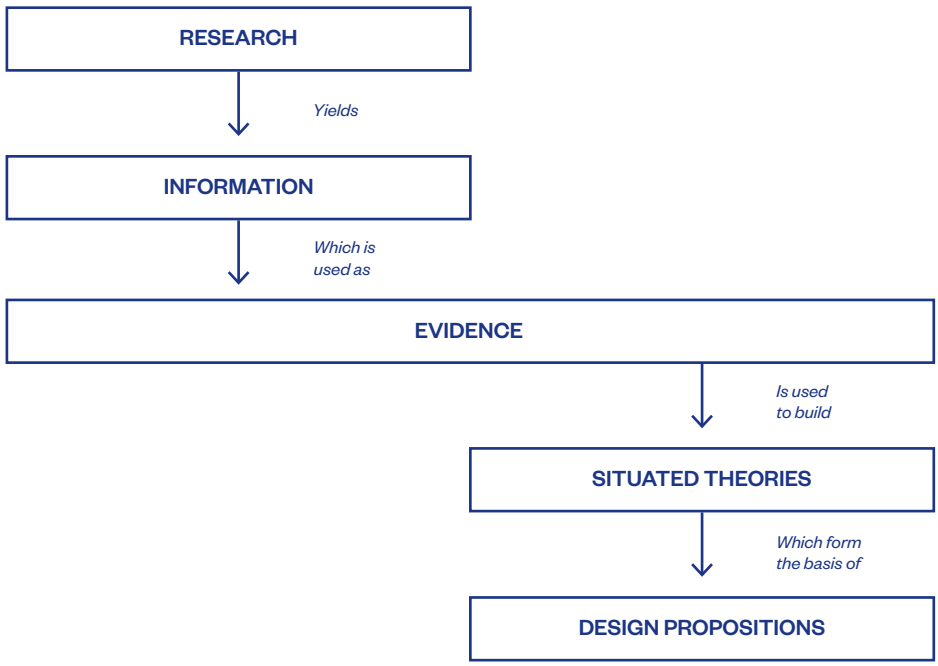

FIGURE 4 A diagram that visually illustrates how information becomes evidence, situated theories and design propositions. Adapted from Miller and Rudnick (2012).

2.3.1 Categories of Evidence

Before different concepts of evidence-based practices are introduced, it is important first to identify the different categories and concepts of evidence.

It is contended, that in sciences, it is basic to an idea of evidence that evidence for a hypothesis constitutes a justifiable reason to trust that the hypothesis is true.[89] It has also been argued that any account of evidence must be taken against the background of what we want an account of evidence for. As Peter Achinstein (2001) has expressed when the claim is made that a hypothesis is true, that something is evidence, what exactly is alleged? According to Achinstein, there is no single, commonly held definition of evidence but several definitions that are used across sciences.[90]

90 Achinstein 2001

TABLE 2 illustrates Achinstein's four categories of evidence that try to explain the relationship between hypothesis and evidence. These categories are *epistemic situation evidence (or ES evidence), subjective evidence, veridical evidence* and *potential evidence*.[91] Achinstein's categories may seem rather abstract, but they show that the definition of evidence is both complex and multidimensional, yet critical for understanding evidence-based practices (introduced in **CHAPTER 2.4**). Some scholars like Reiss (2014) have contended that the issue with these categories is that they are philosophical ideas that have little to do with viable reality.[92] Achinstein thinks of evidence as a '*good reason to believe*', and does not believe that one needs a different category of relevant or supporting evidence, e.g. to support arguments or decisions.[93]

91 Achinstein 2001; Stanley 2004

92 Reiss 2014

93 Achinstein 2001

	EPISTEMIC SITUATION EVIDENCE	SUBJECTIVE EVIDENCE	VERIDICAL EVIDENCE	POTENTIAL EVIDENCE
Description	Evidence related to particular epistemic (cultural, historical or knowledge) context.	Evidence based on or influenced by personal feelings, tastes or opinions.	True or truthful evidence that transcends situations and beliefs.	Evidence stating something is possible or probable – related to experience, present and future.
Requirement	That the inquirer could construct or maintain hypothesis based on the evidence within the limitations of their context.	That the links between evidence and hypothesis are held to be true and consistent by the inquirer(s).	The data supporting evidence needs to be objective, although not necessarily complete (conclusive).	The data supporting evidence must be objective and rigorous and are understood not to be conclusive.
Limitation	The belief is not challenged by ideas from beyond the epistemic context.	It is not necessary for any empirical elements to come into this inquiry.	Both evidence and hypothesis need to be true (very hard to establish).	Hypothesis may be false even when there is a good evidence to support it.
Link to beliefs	The inquirer was justified in believing hypothesis on this evidence, in context.	The inquirer believes that evidence is evidence for hypothesis, that hypothesis is true, evidence does not have to be empirically true, provided that it is believed.	Evidence is evidence for hypothesis and provides a good reason to believe hypothesis, since both evidence and hypothesis are true.	Evidence is evidence for hypothesis and provides a good reason to believe hypothesis until other evidence emerges to challenge.

TABLE 2 Peter Achinstein's (2001) four categories of evidence. Achinstein defines Potential evidence as the central category in terms of which defines three others.

Sustar and Feast (2016) suggest that a more extensive understanding of evidence is critical since argumentative proofs for example thinking from analogy or metaphors are typical for the design practice. The authors propose a new, broader model of evidence and includes '*not only empirical evidence […], but other kinds of evidence such as proofs that support argumentation*'.[94] The first category to include in this new model is *empirical* evidence, that can be defined as the knowledge acquired using the senses, especially by observation, perception and experimentation. The second category is argumentative evidence, that is evidence that supports reasoned argumentation, and is concerned with dialectical and rhetorical perspectives that are closely related to design practice.[95]

94 Sustar & Feast 2016; Ball & Christensen 2009; Sevaldson 2011

95 Sustar & Feast 2016; Buchanan 2001 & 2015; Feast 2012 & 2015

2.3.2 Evidence and Biases

As Kvernbekk (2011) has stated, the principle behind evidential ranking is *trustworthiness*. Any evidence should be trustworthy, reliable, or even true in order to do its job. Evidence is one thing, but high-quality, good, strong and trustworthy evidence is another. Acting on poor evidence may cause more harm than good in any situation, and it cannot prove the truth-value of our claims.[96] This is why good quality evidence is an important part of the scientific process. Ohala (1986) contends that the role of evidence is not to lead us to truth, but rather identify its purpose.[97] According to Voutier (2014), bias is '*any influence, or a systematic error, in the conduct of the study that affects the outcome.*' [98]

96 Kvernbekk 2008

97 Ohala 1986

98 Voutier 2014

Making of decisions is essentially an activity that is related to cognitive thinking and it is seen as a result of thinking that can pass as either irrational or rational. There are various factors that influence decisions made by people including individualistic characteristics such as personality and experience. Looking at it from the psychological point of view, decisions that people arrive at are in most cases based on a number of needs as well as they are augmented by the preferences of the individuals.[99] From the things that people say to the moves we make every day, our reality and that of business-is included a great many choices, both of all shapes and sizes. How one can come to settle on those choices is the consequence of instinct and examination and, by and large, affected by tendencies that they might possibly know about. Biases that exist in the manner that people get to think and make decisions can be deemed as obstacles to the process of decision making).[100] They are capable of distorting and disrupting the objective and factual thinking of issues by bringing in influence on the process of decision making and this is different from the decision. People are mostly unaware of the biases that they have that can affect their judgment.

99 Kvernbekk 2011; Ohala 1986

100 Kvernbekk 2011

The extensive list of around 170 cognitive biases listed on Wikipedia is indicative of the the great amount of research done over the last 50 years, but also indicative of the fallibility of our reasoning.[101] Although the list of biases is long, each type of bias can affect research at any stage of the process, thus affecting the reliability and validity of the research. Some examples of the biases that are rather common include *confirmation bias*, *anchoring*, *overconfidence bias* as well as *the halo effect*.[102]

101 Wikipedia 2017

102 Voutier 2014

- **CONFIRMATION BIAS** is evident when one favours information that affirms already existing beliefs or biases, or settles on a choice as to past evidence.

- **ANCHORING BIAS** is the over-dependence on initial pieces of information including experience that are relied on coming up with subsequent judgment.

- **HALO EFFECT** is is the perception that one has that when another person is so good in a given area, then they are obviously performing well in other areas as well.

- **OVERCONFIDENCE BIAS** is when one overestimates the dependability of the decisions that they are making. Overconfidence bias can include the certainty of the feelings that one has towards their own abilities, performances as well as chances of success.'[103]

103 Voutier 2014

Amos Tversky and Daniel Kahneman (1974; 1983) have uncovered the hidden biases of the human mind in their groundbreaking work, that earned them the Nobel Prize, and paved the way for the rise of the field of behavioural economics. Tversky and Kahneman argue that human decision making occurs in a natural, non-rational way that contrasts essentially from rational models, and, demonstrate that individuals use heuristics to assess the potential value of losses and gains as opposed to the outcome of a decision.

TABLE 3 illustrates the five most prominent biases that can prompt to a poor decision-making, particularly in circumstances including inadequate information, according to Tversky and Kahneman.[104] Also, **TABLE 4** illustrates several common biases that can threaten the validity of studies.

104 Tversky & Kahneman 1974

In the explanation of research confirmation, information that have been gathered in a particular confined review are commonly inspected. In any case, essential inclinations may go before the review outline. A review might misdirect, pointless, or even unsafe, despite the fact that it is by all accounts consummately planned, led, broke down, and announced. A few predispositions relate to setting the more extensive research motivation and incorporate poor logical importance, or inability to consider earlier confirmation, non-thought of earlier proof, one-sided thought of earlier proof, or thought of one-sided earlier proof.[105]

105 Ioannidis 2008

REPRESENTATIVENESS BIAS	The bias results from incorrect associations and occurs when we overestimate what we are familiar with, or underestimate things which we are not familiar or that we do not remember as well.
AVAILABILITY BIAS	The bias operates on the idea that in the event that we can consider it, it must be important. When we can promptly recollect past occurrences of an event, we tend to overestimate the probability that such an event will happen again.
ANCHORING	Anchoring happens in probability assessment as well as when considering new ideas. It refers to the human mind getting fixated on an idea or a number because the person has heard it recently.
OVERCONFIDENCE	The bias affects us all some time or another. Overconfidence is the unwillingness to consider others' opinions or believe data because we overestimate our likelihood of being right. This bias is so ingrained in most decision makers' thinking that it is hard to even provide true confidence intervals even when there is no incentive in being correct.
MOTIVATIONAL	Many times when reporting probabilities of uncertain events decision makers inflate or decrease the values based on what suits them the most.

TABLE 3 Examples of personal biases that affect our decision making. Adapted from Tversky and Kahneman (1974).

SELECTION BIAS	How subjects were chosen to be studies.
ALLOCATION BIAS	How the subjects were assembled into groups.
ATTRITION BIAS	Accounting for subjects at the close of the study.
CONFOUNDING	Other issues present that effect the intervention and outcome being studied.
DETECTION BIAS	The blinding of assessors to which result comes from what group aims to reduce this.
DATA COLLECTION	Were valid and reliable instruments used to assess outcomes?
DATA COLLECTION	Was the sample size sufficient to detect an effect?

TABLE 4 More examples of cognitive biases that affect our decision making and threaten the validity of studies. Based on Voutier (2014).

2.4
Evidence-based Practices

Evidence-based practices is a broad term that is progressively used to describe principles and techniques that spare us from depending on blind faith or anecdotes. As a rule, evidence-based practices are used to create evidence – knowledge that is used to support a perception, claim, hypothesis or decision. According to Mullen (2002), evidence-based practice can be considered any practice, program or policy, that has been set up as viable and effective through rigorous testing and scientific evaluation as indicated by some set of explicit criteria.[106]

106 Mullen 2002

Evidence-based practices originate from the clinical and medical field, and as the name suggest, the overall goal of these practices is to find the highest level of research evidence and use that evidence to make decisions. In healthcare, evidence-based practice is defined as the clear, careful and wise use of best available evidence in making decisions (about the care of individual patients) and a process of making decisions in which conclusions are made using best evidence on a case-by-case basis.[107] **TABLE 5** further illustrates the advantages of evidence-based practices in the context of healthcare.[108]

107 Sackett et al 1996; Mullen 2004

108 Ross 2012

PATIENTS	PRACTITIONERS	ORGANISATION
Reduces the amount of time wasted on wrong or redundant targets.	Professional empowerment through enhanced knowledge.	Enhance quality of service delivery as practitioners can draw upon a variety of options.
Increased consistency and improved service delivery.	Increased personal and professional confidence in problem solving as practitioners adopt a critical approach.	Enhanced confidence in the workforce as decision making is reflected in enhanced care outcomes.
Increased confidence in service providers.	Increased quality of care through patient satisfaction and positive healthcare outcomes.	Reduction in complaints and litigation.
Increased value for money.	Protection against litigation through rationales for action.	Observable commitment to clinical governance.
Reduced variation of services.	Ability to support actions with credible sources or scientific information.	Increased cost effectiveness and value for money.
Evidence can be used to support the need for additional resources.	Appraise options and interventions.	Evidence for the allocation of resources.

TABLE 5 Advantages of evidence-based practices in the context of healthcare. Adapted from Ross (2012).

However, Titchen and Higgs (2001) have proposed that this type of narrow view of evidence-based practices, in which research evidence is the main evidence that matters, should be stayed away from. What counts as evidence varies across disciplines, and some scholars suggest that rather finding the '*highest*' level of evidence, the focus should be in the search for the relevant evidence.[109] Based on the diverse ideas and concepts of evidence found in healthcare literature, Upshur et al (2001) propose a holistic model of four distinct, however related categories of evidence (**TABLE 6**). The model means to legitimise evidence got from qualitative studies and sets it on an equivalent balance with different types of research, expanding the scope of admissible evidence in healthcare decision making.

109 Taylor & Savin-Baden 2001; Upshur et al 2001

In this model evidence can be comprehended as an intervention between the setting of its use and technique for its production. Upshur et al (2001) explain that an important point in this model is that evidence in healthcare is neither exclusively quantitative or qualitative, but is an interaction of both. The model features a wide range of research methods, and describes how there is no single method that would fit all possible circumstances.[110] While healthcare research often looks for empirical data of '*what works*', some scholars argue that a mix of different types of knowledge is necessary to ensure efficient knowledge management between explicit and tacit knowledge forms.[111] Brechin and Sidell (2000) highlight three diverse methods for knowing – observational, hypothetical and experiential.[112] The first is the most explicit type of knowing, which is frequently based on quantitative or qualitative research study. The second way uses distinctive theoretical systems for contemplating a specific issue, sometimes educated by research, however regularly inferred intuitive and informal ways. Lastly, experiential knowing is the craft or tacit knowledge that is built up over some years of practice experience.

110 Upshur et al 2001

111 Brechin & Sidell 2000; Glasby et al 2007

112 Brechin & Sidell 2000

	ILLUSTRATIONS	EVIDENCE TYPE	REASONING	DISCIPLINES
QUALITATIVE (PERSONAL)	Attitudes Perceptions Signs and symptoms	Concrete Particular Historical	Narrative	Nursing Clinical medicine Ethnography Humanities
QUALITATIVE (GENERAL)	Policies Consensus statements Community and social goals	Historical Social	Narrative	Administration Social sciences Epidemiology
QUANTITATIVE (GENERAL)	Traditional evidence Hierarchy	General Mathematical	Quantitative	Clinical epidemiology Bench sciences Statistics
QUANTITATIVE (PERSONAL)	Bayesian Decision theory Quality of life	Particular Mathematical	Quantitative	Economics Political science Statistics

TABLE 6 The model of four categories of evidence serves to incorporate the epistemologies found in the various disciplines involved in healthcare. Adapted from Upshur et al (2001).

2.4.1 Evidence-based Practices in Policy Making

In the context of policy making, evidence-based practices such as field experiments and behavioural interventions present new opportunities for politicians, public servants, government bodies, and official entities to increase impact while reducing inefficiency. These methodologies make another grounds upon which to limit tendencies towards error, and expand the variety of conceivable actions and make opportunities for new and more effective solutions accordingly.[113] Great quality policy making relies on high quality information that is derived from an assortment of sources, for example, expert knowledge, existing and new research, stakeholder consultation and evaluation of past policies.[114]

113 Miller & Rudnick 2012

114 Cabinet Office 1999b

'*What matters is what works*' was one of former British Prime Minister Tony Blair's most used mantras, such that it became a catchphrase in his New Labour government. However, this guideline was not followed in practice, and evidence was often routinely ignored in policy-making.[115] Predictably, Labour lost the 2010 election, but the foundation for evidence-based practices were set in place. Upon becoming Prime Minister leading a Conservative-Liberal Democrat coalition, David Cameron established the Behavioural Insights Team (BIT) within the influential Cabinet Office. The goal of the BIT was to work across all government departments, drawing on the latest academic research in the fields of behavioural economics and psychology. These approaches are typically designed to '*nudge*' behaviour – an evidence-based strategy that policymakers believe will promote individual and societal interests, while also preserving individual freedom to choose.[116]

115 Rentoul 2013; Wright 2014

116 Rutter 2015

'*Nudge*' became a mainstream term in public discourse with the publication of a book called *Nudge: Improving Decisions About Health, Wealth and Happiness* by Richard Thaler and Cass Sunstein. The term is often used to describe a low-cost and accessible change, inspired by behavioural

insights, that accelerates adoption of expected behaviour.[117] Until recently, the theoretical foundations of '*nudge*' theory rest primarily on findings from laboratory experiments – described as the '*emerging science of choice*' – conducted by psychologists concerned with understanding human decision-making.[118] From 2011 to 2012, BIT's team of economists, behavioural psychologists and government staffers put '*nudge*' theory into practice. It backed up its work by rigorous use of field experiments – randomised controlled trials (RCTs) – and demonstrated that it is possible for public services to adopt a '*test, learn, adapt*' approach.[119] In some fields, particularly science medicine, RCTs are regarded as the '*gold standard*' of quantitative research.[120] As illustrated in **FIGURE 5**, RCTs are compelling since they lessen bias by using a randomly assigned control group, to which the results of the interventions are compared and analysed.[121] Numerous experts claim that observing such a non-intervention group is the only way to know whether the policy being tested is due to a change is behaviour.

117 Thaler & Sunstein 2009

118 Yeung 2012

119 Behavioural Insights Team 2012

120 Castillo & Wagner 2013

121 SEE Platform 2013; Edovald & Firpo 2016

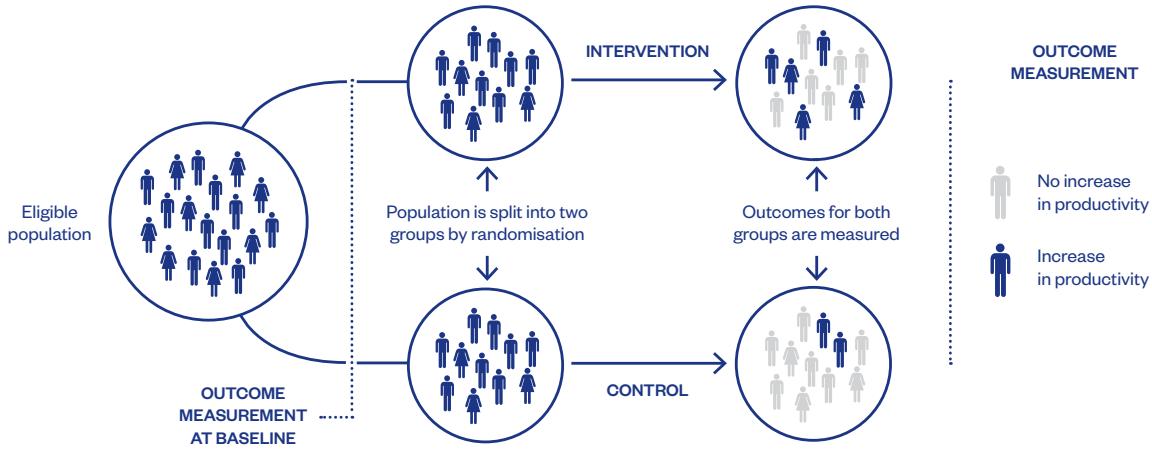

FIGURE 5 An example of a simple randomised design, highlighting the key principles of RCTs. Adapted from SEE Platform (2013) and Edovald and Firpo (2016).

Even though RCTs seem to be increasingly popular especially in the UK, there is also a lot of criticism towards them. Some critics claim that RCTs are full of problems and complexity. One more criticism is that RCTs are narrow in their scope and do not provide solutions to complex problems, and also that is not ethical to purposefully conduct an intervention with a random sample of people.[122]

122 Blume 2016

Despite the criticism, BIT has gained strong popularity and momentum as a model for applying knowledge about human behaviour to shaping policy outcomes. '*What matters is what works*' has become an important approach for not only judging the success of public services, but for also identifying what works, for whom and under what conditions. BIT has proved that evidence-based practices such as behavioural interventions and field experiments can assist governments to pick the best strategy, to spot bias and along these lines to save time and money.

While BIT is not a design organisation as such, they have been paving the way for more agile ways of working, including systematic use of behavioural interventions, field experiments that integrate behavioural evidence into the policy-making process and guiding the development of public services in areas such as tax forms and employment services.[123] While the RCTs provide the strong evidence base to support opinions before implementation, it is not clear how such powerful tools and rigorous evidence can be efficiently used in the public service design process, or is it needed in the first place. Recognising the importance of evidence is essential to the practice of service design, as well as understanding how evidence is collected and how data is interpreted and represented. The most reliable evidence is usually generated through data collection and analytic procedures, since it is a process that is objective, systematic and open to improvement.[124]

123 Bason & Schneider 2015; Mimica 2013

124 Davies & Nutley 2008

2.4.2 Assessing the Quality of Evidence

It has been argued that evidence quality depends on '*What we want to know?*', '*Why we want to know it?*' and '*How we envisage that evidence being used?*'.[125] Evidential quality can be evaluated in view of the source type, and a hierarchical system of grouping evidence is regularly thought to be an imperative part of evidence-based practices. Despite the fact that there is no single, universally-accepted hierarchy of evidence, they can likewise be debated concerning the questions of '*what constitutes evidence?*', or '*what should count as evidence?*'.[126] Drawing from the social sciences and the field of medicine, this hierarchy is better known as the '*levels of evidence*', originally described in a report by the Canadian Task Force on the Periodic Health Examination in 1979 and further examined by Sackett (1989).[127]

125 Nutley et al 2013

126 Cook & Gorard 2007

127 CTFPHC 1979; Sackett 1989

In these hierarchical systems, research design has been used as the major marker of the strength of evidence.[128] As can be seen in **FIGURE 6**, the levels of evidence ranks different types of research designs and rates or grades them from highest to lowest according to the probability of bias. According to Voutier (2014), the hierarchical system is not an absolute measure of evidence, but it is fairly a coherent approach to exhibit the contrasting qualities of various studies.[129] As can be seen in **FIGURE 7** and **FIGURE 8**, these hierarchies tend to rank more sophisticated research designs and generalisable studies on top and user opinions and single case studies near the bottom.[130] The criticism toward these hierarchies is that they tend to put excessively emphasis on research design than on basic evaluation of how that design was executed, and how it fits with different studies about a similar issue. Hierarchies tend to undervalue the contributions of other research perspectives, and also expert or professional judgement. Another criticism is also that certain research designs may not be sufficient in addressing cause and impact of a situation.[131]

128 Nutley et al 2013

129 Voutier 2014

130 USPSTF 1989; Burns et al 2011; Daly et al 2007

131 Kvernbekk 2008; Oancea & Pring 2009; Pawson 2012; Nutley et al 2013; Burns et al 2011

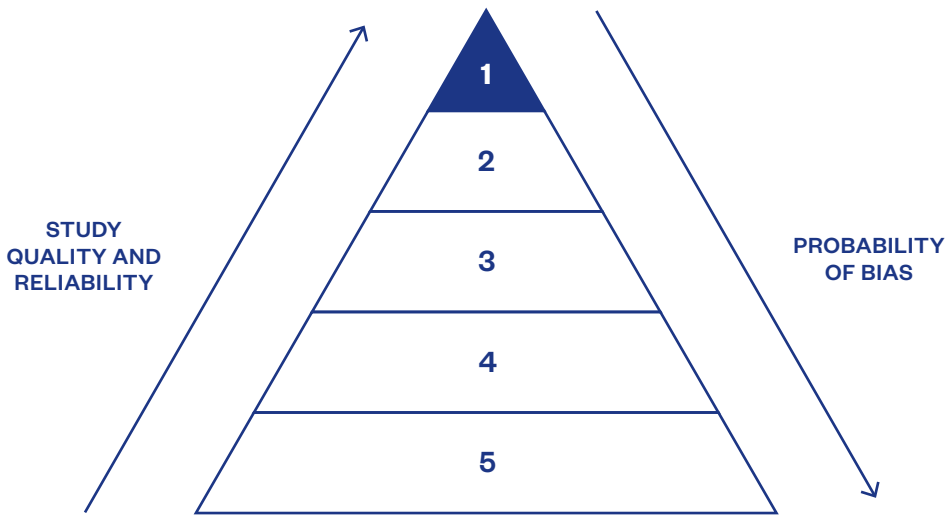

FIGURE 6 An example of the '*levels of evidence*' that demonstrates the strengths of different studies and research designs, and ranks them according to the probability of bias. On the top of the pyramid are the most reliable studies that are designed to be unbiased, and thus have less risk of systematic errors. Adapted from Voutier (2014).

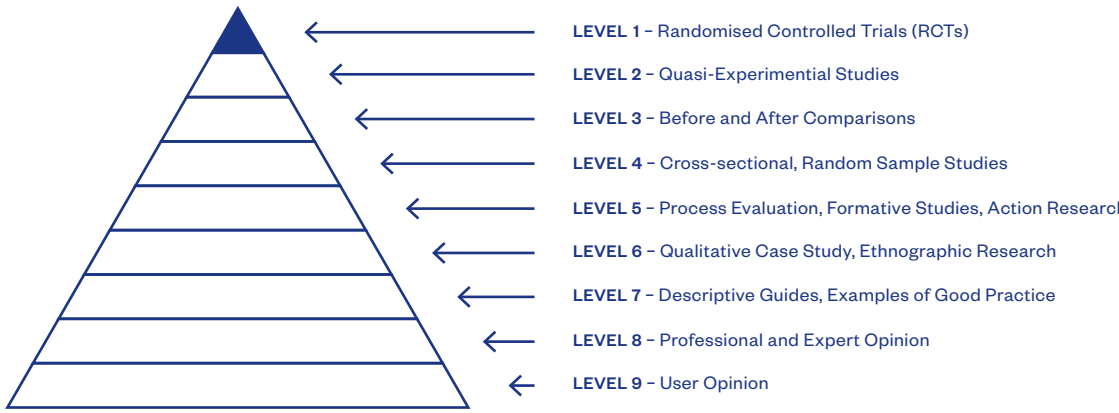

FIGURE 7 An example of the '*levels of evidence*' that represents the scientific hierarchical system of classifying evidence. This pyramid favors quantitative research designs such as randomised controlled trials (RCT) at the highest level and leaves professional, expert and user opinions at the lowest levels. Adapted from multiple sources, including USPSTF (1989) and Burns et al (2011).

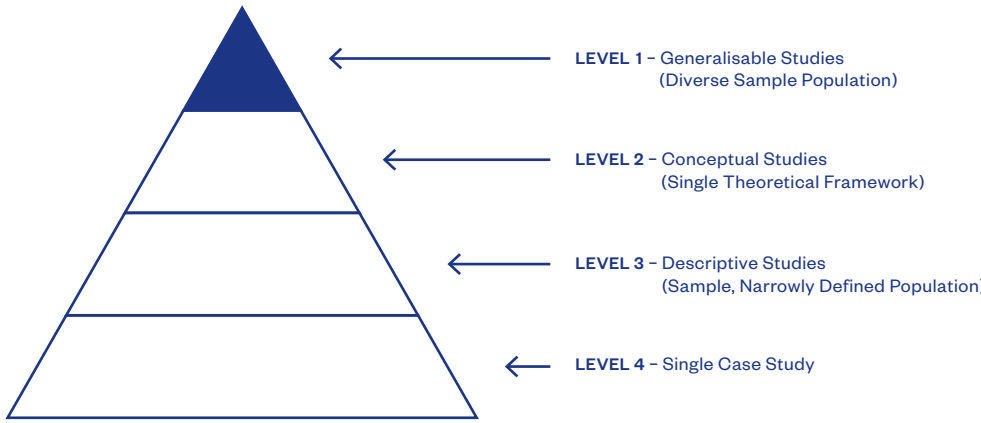

FIGURE 8 An example of the '*levels of evidence*' that represents the hierarchy of evidence for qualitative studies. Since qualitative studies provide different kind of evidence than quantitative research designs, this hierarchy favors generalisable studies over single case studies. Adapted from Daly et al. (2007).

This study has an emphasis on evidence that is supported by qualitative research instead of simply scientific methods that are validated through laboratory studies. In the past, there has been a great debate debate between quantitative research and qualitative research.[132] Some differences between these two traditions in research concern the breadth and depth of the study. Quantitative research provides more rigour and insights into the causes and effects and differences across contexts by classifying, counting and constructing statistical models – which is considered superior when it comes to reporting impact. Qualitative research aims to observe a phenomena with attention to detail, context and nuances – giving us a better understanding into why and how things happen.[133]

132 Bryman 1984

133 Clinks 2017; Patton 2002

In sciences, qualitative research is regarded as less valid and reliable than quantitative research. Some scholars argue that this is because of the fact that it relies upon the researcher's own subjective interpretation of the data, fewer service users are consulted through the research, research participants are selected based on subjective evaluations, and questions are not asked or analysed in a consistent way.[134] However, it is also mentioned that qualitative research can be as credible as quantitative research and should not be dismissed.[135] Therefore, evidence is a standout amongst the most significant figures giving a justification to qualitatively derived research findings.[136] There is no one approach that can provide an equal evidence but approaches must be suitable for the situation or challenge. Therefore, '*What counts as good evidence?*' varies considerably depending on the context.[137] These two approaches have their own uses, and good evaluation combines both research methods. The method in which data is examined in qualitative research dependably relies on the research question, the way the data is collected and the objectives of the study.

134 Horsburgh 2002

135 HM Treasury 2011

136 Miller & Fredericks 2003

137 Nutley et al 2013

As the **TABLE 7** indicates, qualitative research emphasises words and meaning, that we see things from the perspective of the participants, research is practical and close to the social reality.

Qualitative research helps to generate theory rather than being guided by it, it takes the social reality as a process, it is unstructured and emphasises the importance of understanding the context in which events and outcomes occur. Therefore qualitative research offers data that is detailed and rich in depth; further it is micro and has a focus on natural settings rather than controlled or artificial settings.

QUANTITATIVE RESEARCH	QUALITATIVE RESEARCH
Numbers (mathematical)	Words and meaning (descriptive)
Point of view of the researcher	Points of view of participants
Researcher distinct	Researcher close
Theory testing	Theory emergent
Static	Process
Structured	Unstructured
Generalising	Context understanding
Hard reliable data	Rich in depth
Macro-level	Micro-level
Behaviour	Meaning
Controlled or artificial settings	Natural settings

TABLE 7 A comparison outlining the differences between quantitative and qualitative researches. Adapted from Bryman (2008).

2.4.3 Standards of Evidence

Another consideration that should be taken into account to gain high quality evidence is the '*standards of evidence*' framework developed by Nesta. Inspired and influenced by the 'levels of evidence' described in the previous chapter, '*standards of evidence*' were developed in 2012 to encourage the evidence ecosystem in London and the UK. Nesta wanted to better determine how products or services that they fund lead to positive impact on social ventures.[138] Unlike '*levels of evidence*', '*standards of evidence*' doesn't focus on particular types of research methods or particular types of data. Instead, the framework recognises the most promising innovations and continuing to generate evidence to guarantee they are working.

138 Ní Ógáin (2015)

As can be seen in **FIGURE 9**, and further explained in **TABLE 6**, Nesta's framework has five levels starting from the objectives of the research, followed by data collection, selection of comparison groups, creating replication and scaling and finally proper documentation on procedures and methods established. It has been argued that one should not be confined to one evidence searching method: other approaches which can be tailored towards the situation and research needs are preferred compared to RCTs or hierarchies that has been formerly suggested. Nevertheless, these methods provide a structure where evidences can be validated or used as a guidance to check claims that have been made. Nesta's framework is based on tried and tested methods, and it prevents all evidence to be treated as equal and the effectiveness claims on each evidence can be checked and validated properly.[139]

139 Breckon 2015a

LEVEL	EXPECTATION	HOW THE EVIDENCE CAN BE GENERATED
1	You can give an account of impact. This means providing a logical reason, or set of reasons, for why your intervention could have an impact and why that would be an improvement on the current situation.	You should be able to do this yourself, and draw upon existing data and research from other sources.
2	You are gathering data that shows some change amongst those receiving or using your intervention.	At this stage, data can begin to show effect but it will not evidence direct causality. You could consider such methods as: pre and post-survey evaluation; cohort/ panel study, regular interval surveying.
3	You can demonstrate that your intervention is causing the impact, by showing less impact amongst those who don't receive the product/service.	Robust methods can be considered using a control group (or another well justified method) that begin to isolate the impact of the product/service. Random selection of participants strengthens your evidence at this Level, you need to have a sufficiently large sample at hand (scale is important in this case).
4	You are able to explain why and how your intervention is having the impact you have observed and evidenced so far. An independent evaluation validates the impact. In addition, the intervention can deliver impact at a reasonable cost, suggesting that it could be replicated and purchased in multiple locations.	At this stage, we are looking for a robust independent evaluation that investigates and validates the nature of the impact. This might include endorsement via commercial standards, industry Kitemarks etc. You will need documented standardisation of delivery and processes. You will need data on costs of production and acceptable price points for your (potential) customers.
5	You can show that your intervention could be operated up by someone else, somewhere else and scaled up, whilst continuing to have positive and direct impact on the outcome, and whilst remaining a financially viable proposition.	We expect to see use of methods like multiple replication evaluations; future scenario analysis; fidelity evaluation.

TABLE 8 Nesta's '*standards of evidence*' framework further explained by Puttick and Ludlow (2012). In the framework, Level 1 is the beginning stage, where it is about capturing data, continuing to Level 2, which is about seeing if you can link it to positive change, and moving up to trying to have a comparison group and replication by repeating the evaluation, thus confirming the conclusions. Level 5 is where a product or service has very substantial evidence or effectiveness.

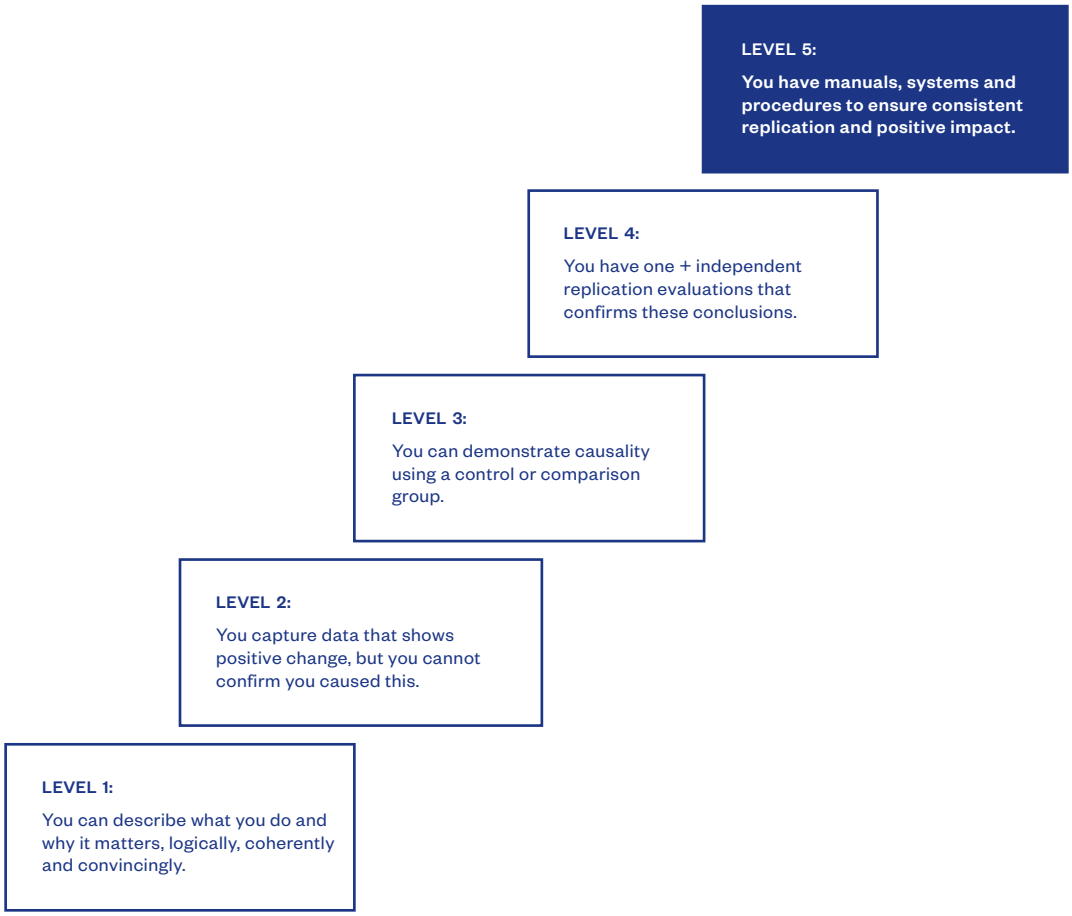

FIGURE 9 Nesta's '*standards of evidence*'. Adapted from Puttick and Ludlow (2012).

The standards developed by Nesta are constantly evolving, and the framework has been further developed by the London-based Project Oracle evidence hub, supporting charitable youth organisations to develop reliable evidence base and more rigorous evaluation systems for their projects.[140] What Project Oracle is trying to achieve, is to measure outcomes in the youth sector consistently, by understanding the starting position ofgiven projects and finding routes for them to improve their evidence base over time.[141] Recent years have seen an increase in the demand for context-appropriate evidence across the social impact sector, and Project Oracle shows that it is possible to evaluate social programmes in line with rigorous standards of evidence, as described in **FIGURE 10.**[142]

140 Project Oracle 2017

141 Ilic & Puttick 2012

142 Project Oracle 2017; Trevithick 2015

Project Oracle is endeavouring to change the attitude of service providers, together with the extensive group of decision-makers and funders, to emphasise the importance of good quality evidence and to create the demand for it.[143] The ecosystem serves the service providers by giving more confidence in their work's impact, so that they are able to provide a better service, and they are able to understand and articulate their offer more clearly. Also commissioners benefit, as they are able to differentiate between work that is innovative, emerging and work that has a strong evidence base. This also helps the commissioners to understand whether projects are really having an impact. The evidence ecosystem also serves academics, that helps them to bring research insights closer to practice.[144]

143 Ilic & Puttic 2012

144 Bediko 2015; Project Oracle 2017

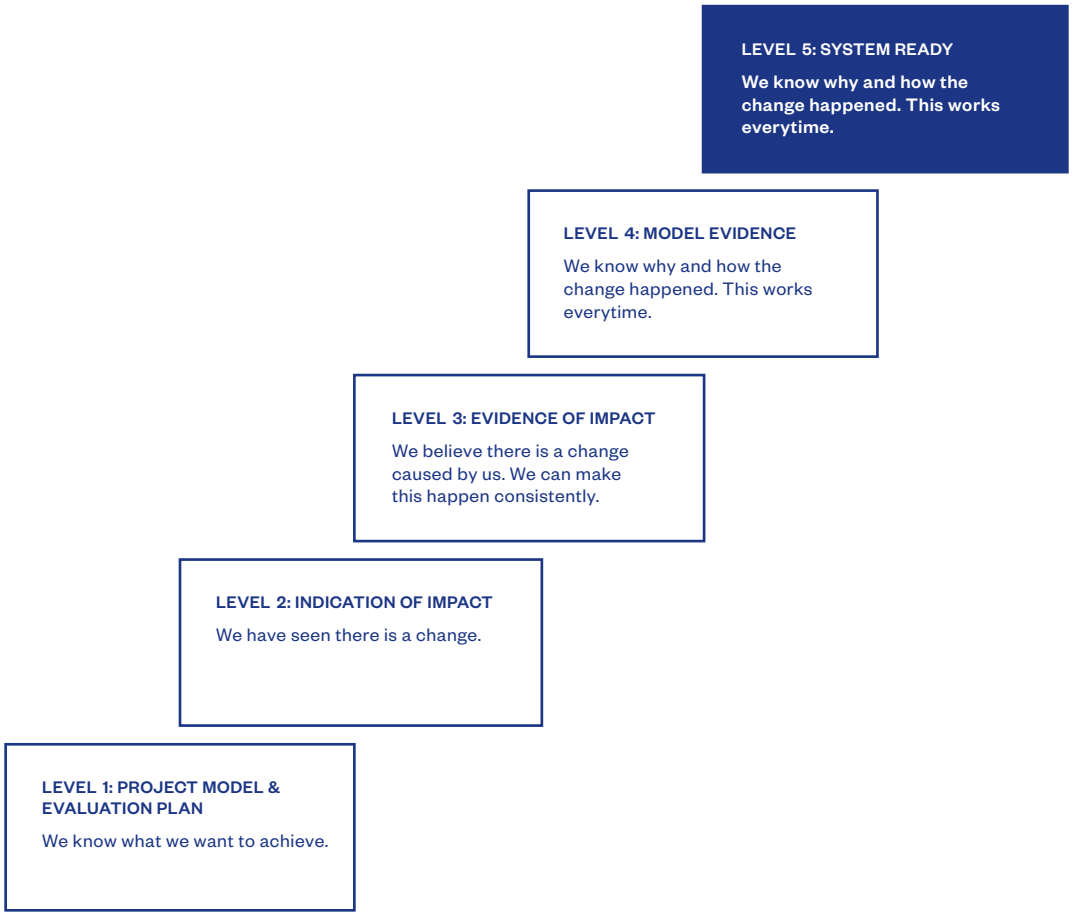

FIGURE 10 Project Oracle's '*standards of evidence*'. Adapted from Project Oracle (2017). The framework is part of Project Oracle's evidence ecosystem in the UK, furthermore, it is designed to address the issues of different social sector organisations and also funders, in a way that supports standardised approach and the use of shared language.

2.4.4 Validity and Reliability

Validity is often referred to as *research validity* within the context of judging the quality of a research. According to Moskal et al. (2002), validity is the extent to which the evidence backs that the interpretation of the data is accurate and the degree to which analysis used is suitable.[145]

145 Morgan et al 2000; Moskal et al 2002

Unlike quantitative research, qualitative research does not have widely used or generally accepted evaluation criteria or guidelines for validation.[146] The issue of validation in qualitative research is somehow contentious and vague.[147] There has been some debate whether the same set of criteria used in qualitative studies can be connected to assessing qualitative research.[148] It has been proposed that the criteria for qualitative research ought to be distinguished from quantitative research, by calling validation something other than *validity* and *reliability*. In qualitative research, consistency and dependability of data and analysis are conceptually similar to *reliability* in quantitative research.[149] Lincoln and Guba (1985) have suggested a process called *inquiry audit* to measure dependability of qualitative data and consistency. They argued that since reliability is a fundamental condition for validity or legitimacy, showing validity in qualitative research is enough to build up reliability. Validity, in the context of a qualitative study, is characterised by the degree to which data are plausible, credible, and trustworthy, and thus can be defended when challenged.

146 Kirk & Miller 1986; Lee & Hubona 2009

147 Maxwell 1992; Ridenour & Newman 2008

148 Lincoln & Guba 2005; Maxwell 1992; Stenbacka 2001

149 Lincoln & Guba 1985; Patton 2001

Maxwell (1992) has suggested three types of validity in qualitative research; 1) *descriptive validity* (the accuracy of what is reported by the researchers, for example events, objects, behaviours and settings); 2) *interpretive validity* (the accuracy of interpreting what is happening in the minds of the participants and how much the participants' perspectives, thoughts, feelings, intentions, and experiences are accurately understood by the researchers); and 3) *theoretical*

validity (the extent to which the theoretical explanation developed fits the data and is credible and defensible.[150]

150 Maxwell 1992

Despite the fact that these suggestions about validity are broad, different scholars have suggested more specific forms of validity for qualitative studies. For example, Lincoln and Guba (2005) have proposed three criteria for judging the soundness of qualitative research and expressly offered these as an option to more traditional quantitatively oriented criteria. These are; 1) *credibility* (instead of internal validity of quantitative studies); 2) *transferability* (instead of the external validity of quantitative studies); and 3) *confirmability* (instead of statistical conclusion validity in quantitative studies).[151]

151 Lincoln & Guba 2005

3
METHODOLOGY

The aim of this chapter is to outline the research methodology for the empirical part of the study. In an attempt to summarise the research design and methods, the chapter will begin with a description of the research design as well as the methods used to achieve the main objectives. This is then followed by the description of the data collection and data analysis. The chapter ends with a discussion concerning the quality of the research.

3.1
Research Design and Methods

This study was conducted using qualitative research methods, and semi-structured interviews, also known as '*thematic interviews*', were used to collect the important data. This approach was chosen in order to gain information regarding the perceptions and representations of the use of evidence in the context of public service design.

First and foremost, the study has the objective of providing an insight into public service design and its role within this context. The study was designed to gather the necessary know-how for service designers within the public sector, supporting the legitimacy and professionalism of service design. Secondly, it was also important to clarify the knowledge gaps that exist to then enhance the overall understanding of evidence-based practices in service design. To do this, the related challenges and opportunities would be assessed. Finally, the study was also designed to contribute directly to this particular research area, and potentially serving as a platform to be used by others to look into the relationship between public service design and evidence-based practices. Ultimately, the research methodology is two-fold; not only is it based on a review of related literature, it is also an empirical study made up of semi-structured interviews with service designers who are experts with first-hand experience in the service design industry. The empirical data collected from the interviews was complemented and enhanced by the interviews with experts on the topic coming from governmental organisations, innovation charities, academia, and think tanks.

As mentioned previously, qualitative research was used during the study since this has the main aim of gathering an in-depth understanding and better insights into a topic using relatively small samples rather than empirical generalisations. There are multiple different ways to collect qualitative data and different techniques can be used for different purposes, but interviewing was seen as the most appropriate method for this study. Interviewing makes it possible to obtain expert opinions and information about the context, by entering into the informant's perspectives and explore their thoughts, feelings and experiences.[152] With the semi-structured interviews in place, it allowed spontaneous questions to be used and a conversational interviewing style. The interviews were guided by a set of pre-formulated questions, giving some structure and setting the frames in which the conversation was to gain momentum, while allowing for some improvisation.[153] Whenever a semi-structured interview is used, the pre-formulated questions need to be open-ended so the informant can express personal opinions and perspectives on a given topic. When compared to other interview methods, open-ended interview is the most informal method, being completely unstructured, whereas structured interview is highly standardised method, utilising closed questions.[154]

152 Patton 2002

153 Myers 2013

154 Patton 2002

In total, 20 semi-structured interviews took place and these provided the empirical data through recordings and written notes taken at the time of the interviews. The key interviews were recorded and soon after transcribed. The semi-structured style refers to the idea of interviewees being presented with a general theme and topics that will be discussed. With this guide, the themes will be touched upon with each interview and the conversation can gain momentum. With the pre-formulated questions in this study, the interviews could be free whilst following a loose pattern; additionally, there was a flexibility for when different themes arose unexpectedly.

In order to choose the interviewees, judgement sampling was used which is a sampling strategy used widely in qualitative research.[155] This benefits the sample because only information-rich cases would be included.[156] Essentially, this meant that the chosen interviewees were people that could teach and discuss the key topics in the best way in order to contribute to the objectives.

155 Marshall 1996

156 Patton 2002

In this case, a list of potential informants were identified and listed based on the preliminary research and partly on the advice from different advisers who are experts in the fields of service design and design research. For this study, judgement sampling is a justified choice since the study of evidence is still relatively new in the field of service design. For many service designers, the topic is still somewhat obscure and unknown which also makes it unique. Therefore, people would need to be identified and interviewed before deciding who were likely to be able to give insights into the topic.

Finally, service designers working for privately owned service design consultancies and experts with some relation to public sector organisations were targeted for the interview stage. Over time, they became two different groups – primary sources and secondary sources. The interviews with secondary sources were used only for background purposes, and to validate desktop research. Primary sources (*Group A*) would include design practitioners such as service designers, consultants, and design researchers that continue to collaborate with various public sector organisations. Secondary sources (*Group B*) would include experts, researchers and other stakeholders with knowledge on the subject or knowledge on public sector, significant to the research. Public sector organisations are important clients for many service design consultancies. Considering up to half of all sales can come from the public sector, this was an important step. One reason explaining the continually developing use of services provided by service design consultancies is the changing public sector landscape, increased awareness of design-led innovation for public services, and increased awareness and know-how of design in general, which makes design as one of the major drivers of innovation in the public sector.

How many interviews is sufficient for a qualitative research? Myers and Newman (2007) suggests that a bigger issue than having a particular '*sample size*' is making sure that the informants '*represent various voices*'.[157] Also, Myers (2013) proposes that once the point of '*saturation*' has been reached, for instance, if there are no more new insights that are being discovered in the interviews, then it is proper to state that no further interview should be directed.[158]

157 Myers & Newman 2007

158 Myers 2013

3.2 Data Collection

To start the data collection process, the potential interviewees, who were selected through judgement sampling, were approached by email to propose the first interview, as is reflected in **ANNEX 1**. For the most part, the reaction to calls were positive and they seemed happy to co-operate. In total 28 interview requests were sent and 25 positive responses were received. Three interviews were cancelled due to time constraints, but substitutes were not needed since the interviewees represented secondary sources, and not the most important group – service designers. Two interviews were also omitted because they contained practically no material answering the research question, and therefore did not contribute to the findings of this study. Therefore, the final number of the informants was 20, and since all of them represented the groups of interest, the sample is derived purposefully rather than randomly. Wherever possible, the interviews were conducted in face-to-face meetings during field visits but there were also occasions where phone or video calls were more appropriate.

As can be seen in Table 9, the interviewees fell into two groups, primary sources (*Group A*) and secondary sources (*Group B*). The primary sources consisted of service designers and those in the design industry with years of experience working with the public sector. Whilst some of the interviewees had more know-how in tactical operations, others had a strategic way of approaching the subject. The secondary sources were interviewed for background purposes only, and thus, were not included in the research sample. Secondary sources were selected on the basis of their expertise and experience with public sector. These included experts working with the public sector, as well as experts from different research and governmental organisations with knowledge on the subject or knowledge on public sector, significant to the study. Regardless of their approach, they were all experts in their respected field and had something to contribute to the study.

PRIMARY SOURCES (GROUP A)	SECONDARY SOURCES (GROUP B)
Service designers and design practitioners working primarily for privately owned service design consultancies. More specifically those with years of experience working work for, and with, the public and social sector, especially within the government.	Experts, researchers and other stakeholders with knowledge on the subject or knowledge on public sector, working in academia, non-governmental organisations, think tanks, government-related organisations, innovation charities and innovation units.

TABLE 9 In order to shed light on the research question from different angle, the informants could be divided into primary sources (*Group A*) – representing the primary research sample – and secondary sources (*Group B*) – that were interviewed for background purposes only.

While most of the interviewees were from Finland, the informants were chosen to include also other countries in order to gain a global overview and a multicultural perspective. The primary sources included representatives from Sweden and New Zealand, while the secondary sources included representatives from Denmark, United Kingdom and the United States.

Although this study does not focus on the specific cultural differences among these individual countries, it is good to point out that these countries vary considerably with regard to the size of their public sector, political institutions and governance. The Nordic countries are modern welfare states with a large public sector, where the systems of education and health service are of a comparably high quality, which is also reflected in high tax rates. As a comparison, the public sector in the United States is relatively small. **TABLE 10** provides a more comprehensive list of the interviewees' countries of origin, and other relevant information. The more specific demographic information was not gathered, as that kind of data was not found relevant to this study.

The interviews were conducted between March 18 and April 30 in 2016, and the majority of them were held in interviewee's offices which allowed them a sense of control and calm. However, there were also two conducted in a local cafe in the city centre of Helsinki which saw some background noise but still allowed the same relaxed atmosphere. Generally, the interviews lasted between 60–90 minutes and were conducted in both Finnish and English – altogether, 26 hours of interviews were recorded. In terms of time, the interviews lasted from 1 to 2 hours, while the majority lasting around an hour. Since there were time constraints in place for the study, only one round of interviews occurred but multiple people were interviewed from some organisations in order to enrich the data and broaden the perspectives. In most of the cases, these varying opinions and viewpoints proved to be an absolute necessity because of the broad nature of public sector and public services. **ANNEX 2** provides a list of example interview questions, sub-questions and probes for the the primary sources.

After the 20 interviews, the collection of data was finished because the data was well-saturated. In terms of sample size, there are no rules as such for qualitative research but the purpose of the study must be considered. Because data saturation can occur, sometimes limiting the number of interviews can yield stronger results – saturation is where the data has reached its maximum effectiveness and further interviews will not bring anything new to the table.[159] As an iterative data collection process, analysis of the data guided the way both during and after the collection; in terms of the interviews, these underwent sequential analysis throughout the data collection process. When data related to relevant themes, these were highlighted and examined to find a clear and coherent understanding of the definition and the role of evidence. Also, it was additionally essential to note how evidence was being used amid the public service design process. To reveal and identify emergent themes, a preliminary thematic analysis was led on the data from the interviews and these topics were then used as a guide for further questions in the following interview. Consequently, the questions would be classified as new and relevant issues that emerged from the early interviews.

159 Patton 2002

The actual wording was captured from the key interviews using recording devices and notes were also made. As well as showing emphasis on certain topics, they allowed the questions to be formed for the latter half of the interview. Wherever the discussion was being recorded, the need for making verbatim notes vanished. Therefore, it was possible to engage with the interviewee fully which led to in-depth conversations. After the interview had finished, the notes and recordings were checked for quality and all relevant details were noted. In terms of debriefing, this document was created soon after the interview and the transcription for key interviews were in both Finnish and English – the same can be said for the citations. Citations were later translated in English for this study, and in the translation process, the meaning of the original citations were given special attention.

TABLE 10 Interviewee information: primary sources (*Group A*) and secondary sources (*Group B*).

PRIMARY SOURCES (GROUP A): INTERVIEWEE INFORMATION

ID	COUNTRY	INDUSTRY	INFORMANT	ORGANISATION SIZE
A1	Finland	Design	Service Designer	51–200 employees
A2	Finland	Design	Service Designer	51–200 employees
A3	Finland	Design	Senior Service Designer	11–50 employees
A4	Finland	IT / Design	Senior Service Designer	51–200 employees
A5	Finland	IT / Design	Service Designer	51–200 employees
A6	Finland	Design	Service Designer	11–50 employees
A7	Sweden	Design	Senior Service Designer	11–50 employees
A8	Sweden	Design	Senior Service Designer	11–50 employees
A9	Sweden	IT / Design	Senior Service Designer	51–200 employees
A10	New Zealand	Design	Design Researcher	1–10 employees

SECONDARY SOURCES (GROUP B): INTERVIEWEE INFORMATION

ID	COUNTRY	INDUSTRY	INFORMANT	ORGANISATION SIZE
B1	Finland	Public Policy	Senior Advisor	> 10 000 employees
B2	Finland	Public Policy	Senior Advisor	201–500 employees
B3	Finland	Academia	Researcher	1001–5000 employees
B4	Finland	Civic & Social	Researcher	11–50 employees
B5	Finland	Civic & Social	Managing Director	1–10 employees
B6	Finland	Public Policy	Assisting Specialist	51–200 employees
B7	Finland	Civic & Social	Development Director	11–50 employees
B8	Denmark	Civic & Social	Executive Director	11–50 employees
B9	United Kingdom	Public Policy	Senior Advisor	11–50 employees
B10	United States	Civic & Social	Managing Director	11–50 employees

3.3
Data Analysis

After the data has been collected, data analysis is required because it aims to make sense of what has been said. By compiling the research, patterns can be found and this starts to build findings. According to Patton (2002), there are many ways data can be analysed but the researcher needs to choose a method that suits the situation and research in question. With this in mind, it is very much the purpose and audience that guide the process. Regardless of which method is chosen, the findings need to be presented with honesty.[160]

160 Patton 2002

With the twenty semi-structured interviews, there were 26 hours of recording, transcripts, debrief documents, and written notes to assess. Since one person was in charge of all interviews as well as completing the transcription process, there was a sense of familiarity throughout and the data could be assessed intensely. After transcribing, the interviews were read through carefully several times to form a basic level of understanding – special attention was paid to challenges and opportunities relating to the role of evidence. In this study the analysis was based on thematic analysis which is widely used within qualitative research. As a useful tool, it allows the data to be broken down into different sections and patterns according to specific themes – a theme represents an important issue about the data relating to the research question and then captures a somewhat patterned response or meaning. In the early stages, themes are yet to be established so qualitative research is normally inductive. Once the themes have been established, it can then be deductive.[161]

161 Braun & Clark 2006; Patton 2002

By using inductive analysis, themes were identified and listed early on and this allows the first step of coding the data into various topics – coding refers to the creation of manageable classifications according to which the data will be grouped. When a code is present, it integrates different instances of the data. Soon after, a coding manual was established to systematically code the data. Essentially, a manual will indicate a name for a theme before then defining its meaning and generating a description of how the theme will occur in the data. When reading the data for a second time, a formal coding procedure was started.[162]

162 Patton 2002

According to the question and theme, the data from the various sources were brought together into an index sheet and the transcriptions were the focus for much of the study. With this in mind, some of the most memorable and applicable quotes were pulled and added to the sheet before then explaining the sources of each. As the transcribed data was verified for accuracy, the themes soon emerged as we coupled this with reviewing notes from the interviews. As the themes were recognised, they could then be colour-coded so that each text was coloured according to its theme. Before the data was completed, several readings occurred until the text relating to different themes were separated so the meaning could be examined. Once again, this was done with the patterns in mind to see which way the data fell. Additionally, differences among various data patterns were also examined. In order to organise large pieces of data and put them into natural relationships, a complete affinity diagram was used as this allowed the issues to be cross-examined. From here, the cause and effect relationships could be found between different themes.

3.4
Quality of the Research

Traditionally, the quality of academic research is evaluated using three criteria – reliability, validity, and generalisability. In quantitative research, these measures are common but qualitative research yields different knowledge and therefore requires different evaluation. It is imperative to note that, in qualitative research, there are no absolute or objective criteria that indicate whether findings are valid and their procedures robust. However, the three measures can still be interpreted from a qualitative perspective.[163]

163 Moisander & Valtonen 2006; Miles & Huberman 1984

Reliability

Describing the degree to which the findings are independent of accidental circumstance, reliability is closely related to the concept of replicability. Essentially, this shows that other researchers could potentially repeat the study and yield the exact same results, claims, and interpretation without having any relation to the original.

With qualitative research, this is a little harder since it is more contextual of time and place. For example, the role of evidence has not been studied in detail and evidence-based practices in service design has not yet matured which means that the research will continue to evolve along with the industry. In the future, research is likely to deal with different challenges and opportunities. Besides, reliability can likewise manage the transparency and straightforwardness of the qualitative research process. In this manner, it is

essential to describe the procedure that shapes data collection and analysis in detail. In order to ensure transparency, this study meant to give a detailed description of the data collection and analysis process. Furthermore, reliability can also deal with the transparency of the qualitative research process. Thus, it is imperative to describe the process that shape data collection, analysis and interpretation in detail. To ensure transparency, this study aimed to provide a detailed description of the data collection, analysis and interpretation process.[164]

Validity

With validity, it refers to the truth or accuracy of the claims made by the study – also seen as how well a study represented the features of a phenomenon that it was attempting to describe. Because of this, the criteria used to assess the nature of a qualitative inquiry depend on the theoretical and methodological points of view initially decided for the study. Since there is no universally-accepted method of evaluating qualitative research, validity is more of an agreement of the interpretations of research judged by the audience.[165]

165 Moisander & Valtonen 2006

Generalisability

Finally, this third variable refers to the degree to which the findings from the study can be connected to a larger population. In qualitative research, this isn't normally the main goal because it tends to deal with smaller samples that do not represent the larger population. Instead, qualitative research looks to expand the understanding of a specific phenomenon in specific content. Therefore, it is perhaps more relevant to discuss transferability which is the idea that readers of the report could establish a connection between the results of the study and their own personal understanding or experiences. Just as this study intends, it becomes imperative to provide a thick description of the research setting.[166]

166 Moisander & Valtonen 2006; Patton 2002

Limitations of the Study

Even though the study reached its aims, there are some unavoidable limitations and this starts with the time. Time was a major limitation, as data had to be available to be analysed in a timely manner to complete the study. In the course of the qualitative nature of the study, the sample involved was relatively small in any case, sufficiently expansive to demonstrate thematic consistency. In order to gain an accurate review of the service design industry, 20 interviews is simply not enough. Considering the other experts and professionals who were also involved in the study, there were only 10 service designers interviewed. With more time available, or with different research methods such as surveys, more comprehensive research could take place.

Informant Confidentiality

This study took after the dominant approach, assuming that informants need secrecy and assuming liability to alter the data to guarantee privacy.[167] Bazeley & Jackson (2013) state that '*if the source is an interview or similar, we would suggest that you substitute pseudonyms – it reduces the risk of breaking confidentiality*'.[168] The underlying assumption that the informants have chosen to be anonymous is not always the case, but some of the informants wished to remain anonymous. Maintaining informant confidentiality while presenting rich, detailed descriptions was one of the challenges for the study, but every effort was made to comply with those wishes.

167 Kaiser 2009

168 Bazeley & Jackson 2013

4

FINDINGS

This chapter is focuses on presenting and analysing the findings of the study. From the transcribed material, there were found many similarities and repeating themes, that helped form the four chapters that are the research result.

4.1
Results

The chapters are grouped in the following logic; the first chapter describes service design in the public sector in general, while the remaining chapters follow a generic service design project chronologically, discussing about the role that evidence plays in public service design.

Service Design in the Public Sector

In the first chapter, the informants discuss about their experiences on the different types of service design projects in the public sector. This chapter examines working within the public sector, the operational environment of the public sector, different tendering procedures, the special characteristics of service design projects in the public sector, the role of the service designer and the importance of the co-operation within these projects. These factors are important to include as they provide first-hand insights for better understanding of the nature and role of evidence in public service design.

Defining the Problem and the Solution

The second chapter is about the defining phases of the problem and the solution, as well as collecting and generating evidence using qualitative methods and service hypotheses.

Developing Solutions Through Experimentation

The third chapter is about experimental developing, as well as collecting and generating evidence through experiments and prototypes, and measuring and evaluating them.

Validation of the Impact

The last chapter is about determining the success in the context of a service design project, as well as the role of evidence in the implementation of the change and validation of the impact.

4.2
Service Design in the Public Sector

In the following, the informants share their experiences about the operational environment in the public sector, starting from the tendering phase, all the way up to the implementation phase. For the purposes of this study, the informants were also asked to share their opinions on what type of service design projects are done in the public sector and what is important in them.

Public sector innovation is not a new idea, but rather generally, the structures of government have been ineffectively suited to handling complex issues that cut across traditional organisational silos and boundaries.[169] The public sector usually lacks the capital, resources, and skills to take promising ideas and thoughts to scale. To respond to big questions, such as how to cut one's carbon footprint or how to make one's cities more liveable and sustainable, a major shift in thinking is required in order to deliver high-impact, insight-driven, and meaningful solutions. However, the amount of regulation and other '*red tape*' imposed by the government can be a major obstacle to public sector innovation.

169 Hartley 2005; Mohr 1969

In the recent years service design collaborative innovation and have become somewhat of buzzwords in the public sector, government and modern politics around the world. In the era of constrained financial resources and increasing public expectations, there is an ever-growing demand for better service delivery in the public sector. Even though the governments are responsible for public services, these services are not provided by the governments alone.[170] More and more privately owned service design consultancies are responding to this

170 Mager 2016b

demand, taking a human-centred approach to addressing different societal problems at a systemic level, formulating and implementing policies, boosting the value of existing services as well as creating value for society through the creation of new innovative services.[171] It can be said that private-public partnerships that create profits for privately owned consultancies and delivers value for money for the general public, fuels economies and societies at its best, and enhances the quality of life of the people, having the best effect on healthcare, education, communications, and transport respectively.[172]

171 Bason 2010b

172 Service Design Network 2016

This study focuses primarily on the role of privately owned service design consultancies, of which many have dedicated an increased amount of resources for the public sector projects. Because the number of service designers working in the public sector is currently either very small or non-existent, many public sector organisations are still in the process of building their own internal service design capacity. These organisations typically buy their service design services mainly from third service providers, such as service design consultancies, whom the service designers interviewed for the purpose of this study, represented. Every one of the informants had experience working within the public sector, which was relevant for the purposes of this study. The informants were chosen in light of their capacities, experience and knowledge that would benefit the study in whole.

While private sector projects, being more business and customer focused, are still a core business for most these consultancies, there has been has been a dramatic increase in the demand for and use of service design in the public sector. A perfect example of this is a service designer (A1), who has a long experience of working as a service designer for a major service design consultancy in Finland. His work is dedicated almost exclusively to public sector projects at the moment, and he is focusing primarily on the development of social and healthcare services. The informant explains how the amount of public sector projects has grown significantly and so has the size of project teams.

> *'The share of public sector projects is growing all the time. I'd say it's about one-fourth of the turnover at the moment. When it comes to actual doing, I'd say third of our employees are more or less involved in public sector projects.'* (A1)

One of the informants (A7), who works as a service designer in a service design consultancy in Sweden, also outlines how his work increasingly involves working with public sector organisations.

> *'I'd say that at the moment the share of public sector projects is about fifty percent of the company's turnover. When we work together with public sector many times, it becomes a long-term relationship, like an account, more or less.'* (A7)

A service designer (A2), who works as a service designer for a service design consultancy in Finland and who has almost a decade of experience working with the public sector, concludes that the amount of work for the public sector has been steadily increasing over the years. Also, commissions have become more interesting, yet vastly more complex.

> *'We have several people who work primarily with the public sector. The number of projects has been increasing all the time, and also the projects have also become much more interesting. In the beginning, they were mostly about workshops and training, a kind of sketchy things. But, at the moment we are involved in a massive project where we will be ideating possibilities to underserved youth with immigration backgrounds, to ensure that they don't drop out of school or workforce. At this point in time, we don't have a clear solution, but the problems are well known. These problems are life-size problems.'* (A2)

4.2.1 The Operational Environment in the Public Sector

In the following, the informants sharing their experiences about the operational environment in the public sector. As described in **CHAPTER 2.1**, the operational environment in the public sector differs significantly from the private sector. When compared to the private sector, public sector is more formalised and more strictly managed due to the fact that it manages public money responsibly and is therefore accountable to the general population to run its operations efficiently.

Due to the special nature of the public sector, public sector officials such as public servants are in a unique position of trust and are accountable for fulfilling their duties with responsibility and diligence. In spite of the fact that there are contrasts between nations, the public sector in democratic and legal states regularly works in an incredible line of hierarchy and accountability that begins with civil servants who are responsible to their hierarchical superiors, and through Ministers, ultimately to the Parliament.[173] This type of bureaucracy is something that makes a qualitative distinction between public sector and private sector. One of the informants (A2), explains that when working with the public sector, one has to take into consideration these special characteristics of the sector.

173 Hodge & Coghill 2004

> *'Depending on the client, service design projects allow you to operate in the public sector like in any other environments. Where in the private sector you can make big things happen just by convincing the CEO – that kind of direction can't be found within the public sector. For example convincing the mayor doesn't do you any good. A branch director or bureau manager can be a more meaningful agent in their own right, but it's possible even convincing them isn't enough, because*

> *there can be a committee to take a final decision, that's leaning politically left or right. And sometimes the means of justifying procedures can be dirty, when there's a lot more at stake than only the solution to the matter at hand.'* (A2)

Due to the special nature of the public sector, public servants are restricted in what they can do and how they work. One informant (A4), stressed the fact that they are too often limited by the assigned responsibilities outlined in their job description.

> *'The public sector has that basic problem that as a public servant, you can't go far beyond your formal job description or branch of government you're part of.'* (A4)

One of the informants (A6), who works as a service designer for a service design consultancy in Finland, argues that there are noteworthy contrasts among the public sector organisations regarding the environment. Some public sector organisations can be exceptionally various leveled where individuals hope to work with obviously established lines of authority, rules and regulations – while some different organisations can be immensely unique.

> *'Countless excuses and obstacles can be manufactured. For example, working in another location, or working in the evenings may not be possible. It's just not possible. Sometimes the collective labor agreements, trade union movements, and such also bring their own impact to this scenario. But those vary, too. Some city agencies can have a lot of flexibility for the personnel, through history and culture. Then there can be another agency where they're really tight, and where all new things are out of the question and no exceptions can be made.'* (A6)

With regards to public sector organisations, the concept of bureaucracy, or '*red tape*', has the connotations of being an inflexible, inefficient and regulation concentrated method for working.[174] One of the informants (A1) explains that bureaucracy is a challenge in the public sector projects.

174 Eskelinen 2002

> *'Bureaucracy is a thing that everyone in the public sector talks about, but then when you ask what it means, you get a whole variety of answers. It can mean the slowness of the decision making, all kinds of reporting, or job descriptions programmed too far. Bureaucracy is clearly a big problem.'* (A1)

Also one of the informants (A2) explained that in the worst-case scenario, bureaucracy can be an obstacle for change and getting agendas approved, slowing the process and productivity.

> *'At its worst, bureaucracy is shifting the responsibility always to somewhere else, up the decision making ladder somewhere up, which means an enormous slowing factor for all actions. People are stripped their ability to think anything on their own. You can't even decide, what kind of a pen to use. In a way, helplessness in organisations is almost tangible. They don't dare or want, or even actually know how to do things differently - when it has been suggested that this is the way things should be done.'* (A2)

Public servants play a central role over the control of the bureaucracy, affecting the administration and enforcement of policy in important ways.[175] However, public servants are often subject to inaccurate stereotypes, such as being dull and faceless government bureaucrats.[176]

175 Johnson & Libecap 1994

176 Donovan 2013

One of the informants (A4), who works as a service designer for a service design consultancy in Finland, outlines that even though these stereotypes are not entirely false, not all public servants are bureaucrats in the real sense of the word.

> *'In the public sector, there can be found the eager developer, who can speak with reason and common sense. But of course, there are also those, who play with the responsibilities of public servants, and do everything 'by the book', which means if something goes wrong, nobody takes the responsibility for it. And when you go to higher level decisions, the decision making process becomes ever more complex when things become politicised. There are a lot of processing going on different issues, power struggles, conflicts and internal tension.'* (A4)

Public servants, who often are also project owners in the public service design projects, are a diverse mix of experts with a wide range of backgrounds, skills, capabilities and responsibilities. Although their job description may be limited to a very specific scope, there are positive examples of public servants who use their expertise beyond bureaucratic boundaries, as an informant (A2) illustrates.

> *'Nowadays there are also public sector leaders who are creative visionaries. They are from a different world. On the contrary, if you go to some smaller town, where officials are selected with political party affiliations and seniority, the approaches can be quite different. There can be someone, who just happens to be in the position, and you might get lucky if they are even a little bit interested in challenging themselves or their organisation.'* (A2)

4.2.2 Tendering Procedures in the Public Sector

In the following, the informants sharing their experiences about the tendering procedures in the public sector. In countries where procurement and tendering systems are in place, the informants explained how public service design projects typically start through the public tendering procedures where a formal and structured invitation to tenders is published in a specialised portal, where providers can submit a bid to supply products or services – in this case service design consulting. One informant (A8) from Sweden describes the procurement process, which is typically regulated by the procurement law.

> *'There is a law in Sweden that says that you have to go through a specific process. They have to open up the call to what they want to achieve, and then companies can apply for this and they have to fulfil certain requirements. This is the process how we do it at all times. Depending if it's a big project of 2-3 years, or a small one of 6 weeks.'* (A8)

In Finland, the invitations for tenders are sent through public procurement system HILMA, where public organisations are obligated to announce all the purchases with the values going beyond the national threshold levels.[177] Similar systems exist worldwide for posting invitations for tenders, such as Tenders Electronic Daily (TED), where all tenders by European governments are published.[178] One of the informants (A3) described that typically these tender invitations lack understanding about what service design can offer to them.

177 Arlander 2014

178 TED 2017

> *'When you have the public sector as a client, they almost always publish an invitation for tenders through HILMA. And typically, they know nothing about the offering of service design. This competitive tendering of public projects is a question for the*

ages, as in how they should be organised - is there enough design knowledge there, that they can do it properly?' (A3)

The informants from Finland and Sweden described, that the formal tendering process is not ideal for buying the service design process as such. One informant (A2) explained, that there is a lack of practical tools and support for the public servants in the tendering phase, and that the current tendering system is not designed for defining the problem or the preconditions in the tendering notices.

'It's easy for the buyer to go to the basic competitive tendering, that's been built for situations, where the one ordering knows very clearly what they want as a result. But this kind of procedure sends the whole thing in the wrong direction, because it's about ordering the end result. It functions really badly when buying a process. And when the buyer sets off to buy a process and realises that they don't know the end result, the whole thing falls apart instead of asking or describing the problem or the preconditions.' (A2)

The informants explained that even though the tendering process is intended to be open and transparent, the competitive dialogue process is under-utilised. However, the requirements in the tendering notes don't always necessary have a natural link to what is being procured. A service designer from Sweden (A8) suggested that it would be beneficial to consistently engage in competitive dialogue with the public organisations to discuss all aspects of the proposed contract, ensuring clearly defined and well-articulated requirements up front.

'We are talking to government departments and different organisations in the public sector, and having a dialogue with them. But I think half year before they start this process of getting applications, they are not allowed to talk to companies. So, they are writing the application themselves, and set vague

requirements. Of course we are helping them and they also want to learn how to do it. Sometimes the invitation for tenders doesn't state explicitly, that they are looking for service design, but we might still apply – and sometimes we also win the bid. And that's how they learn about service design.' (A8)

One informant (A2) also called for more transparency and dialogue in the tendering process. He described how the requirements of the invitations for tenders are sometimes unreasonable, and the issuing organisation doesn't necessarily know how to buy service design. In the worst scenario, this can lead to complaints from the service design consultancies and other service providers.

'The invitation for tenders might for example have a copy of a table of contents from some service design book and a list of everything in it. And it says that if all this could be had with this much money. It shouldn't go like that. The structures of public sector dictates the direction in a way, that it's hard for the individual to do anything. Not too long ago, no-one even bothered to make complaints about the tenders, which makes it even harder for the public servants. If the public servant knows that there will be complaints, they'll be paranoid and try to make sure they can't get caught on anything.' (A2)

More than one informant (A1, A2, A6) mentioned, that there are several ways to select a contractor from bidders in competitive bidding, such as framework agreements. European Commission has defined that '*framework agreement may be awarded by a procedure without a call for competition if the framework agreement has been concluded in accordance with [applicable legislation]*'.[179] What this means in practice, framework agreements are means by which potential suppliers are examined once by the public organisation, and if they meet the prerequisites, they are then viewed as preferred suppliers, thus simplifying and improving the process.[180]

179 European Commission 2005

180 Murray et al 2010

One informant (A6) told, that this type of contracts have become more common nowadays, but currently majority of them give the cheapest purchasing price a clear vantage, thus eliminating those with years of experience in service design, that hold the most potential for addressing the challenges of public sector.

> *'Framework agreements are good in principle, but often public sector organisations don't know how to order service design. The ones that could most likely help with the solving of the problem, stay out of the competition. Often times the providers are selected based purely on offers, and at the end the final choice will be made by the cheapest price. Then they get a mixed bunch of consultants and providers, that wanted to price themselves at half of the market price for one reason or another.'* (A6)

One of the informants (A2) shed some light, that sometimes the lowest price isn't always the deciding factor in the public sector, and that it is possible to go around the formal tendering procedures, by issuing the tender for a fixed price contract rather than sticking with potentially risky tendering process. Whereas formal tendering process applies to more large and complex jobs, fixed-price tenders could be more suitable service design projects that are more agile in nature, providing faster and more actionable insights.

> *'There is a so-called 'French contract', better known as fixed price tenders, where we're told upfront that we have a certain amount of money at our disposal and how would we go about solving this. In this kind of situation these kinds of projects the qualitative work is emphasised in the right way.'* (A2)

The same informant explained also that the fixed-price tenders are reasonable when the scope of work has been a well defined by the organisation:

> *'When a public organisation can determine the price of the service they are buying, the quality of the work stands out in a completely different way. Especially, if their finances are tight and the projects should stay on budget, this kind of fixed price tenders can help to ensure the professional capabilities of the companies offering service design.'* (A2)

The informants illustrated that working with service design requires a shift in mindset for many public sector organisations, moving away from gathering requirements and towards developing hypotheses, learning through experimentation and prototypes. In many cases, it also requires looking at the bigger picture and looking beyond the silos and artificial divisions created within the public sector. Therefore, an understanding about service design and its possibilities is needed by the issuing organisation. One informant (A5) emphasised that understanding about service design and best practices are slowly turning up as service design becomes more widely adopted within the public sector.

> *'One person might have a good understanding about things, but often times the other employees within the organisation don't share the the same understanding. Awareness grows understanding – and the more understanding grows, the more support grows with it.'* (A5)

One of the informants (A7) explains that there has been also positive development in the tendering procedures in Sweden, in terms of formulating the invitation for tenders.

> *'More and more, there has been an evolution in a way that they now explicitly ask for service design, and they want someone with years of experience on service design. It has become more desirable.'* (A7)

4.2.3 The Characteristics of a Public Service Design Project

In the following, the informants talk about what are the different types of service design projects in the public sector that they have been involved in, and how they are creating value together with the decision makers and creating better policies with decreasing financial resources. The informants explained that the broadness of the projects and the focus points can vary, depending on the brief of organisational need. The informants described, that the public sector service design projects can be about concepting services, doing analysis, foreshadowing and strategic work for service development purposes, designing service processes, designing service environments, educating, coaching, capacity building or transformational leadership. A service design project can also be a part of the organisation's own brand communication and marketing development.

There are certain projects and implementations, to which the general public decides to attach greater importance and significance. When assessing the complexity of a potential project, factors such as immediacy of delivery, long-term impact, necessity and economic feasibility each need to be duly addressed. One of the informants from Sweden (A9) explained that the special characteristics of the public sector set clear preconditions and criteria to everything that's done.

> *'The purpose is to create added value to those, who use public services. Added value can be something, that solves some specific problem or need. Or then it can also be something the buying organisation hasn't even realised they're missing.'* (A9)

The informants described how important it is to go in the field and experience the problem yourself, as well as empathise with those who are directly affected by the problem. One informant (A6) explained that in order to understand how a public service might operate, customer insight data is crucial

to not just understanding people as individuals, but also the relationships they have with others. It's not enough to empathise with the real-life concerns of citizens and stakeholders; one must also understand how they are connected together within the larger system.[181]

181 Mazé 2016

> *"In the heart of service design projects, there's typically the accumulation of the customers' and organisation's employees' needs, behaviour and the factors influencing these. When the problem field has been thoroughly mapped out, can this customer insight be used on a strategic level. This can mean defining a more customer centric strategy, steering the operative activities or resourcing within their own organisation, or implementation to fulfil the customers' needs.'* (A6)

Most of the interviewed service designers explained, how public sector projects are often about managing change, and described how the projects tend to move on two or multiple different levels. More practical levels deal with customer insights, that aim to create an emotional connection to the customers through various methods such as interviewing and observing people, while strategic levels often deal with simplifying workflows, clarifying internal concepts and building trust. One of the informants (A7) described these levels as '*touchpoint level*' and '*strategy level*'.

> *'In most of our projects, we work both in the strategic level and touch point level, and we work between those levels. But there is no really big difference there. It's just how we apply and where we apply our processes. If it's both working with customer insights, but also employer insights and employee insights. And we work with the connection between customers and employees, and the overall strategy of the organisation.'* (A7)

Also one of the informants (A8) divided the projects into two different types of categories.

> *'I would say that there are two types of public sector projects. Some of them are more change projects, and we work together as partners and try to accomplish together. Then as well we have shorter projects in which the client just buys our services as a supplier of service design. Shorter projects can last only a few weeks, when we work on a specific issue, for example different parts of the pension system that they want us to focus on. Sometimes they buy one part, sometimes they want a bigger project.'* (A8)

Service design projects are not just about service innovation, and also include training and educating the client organisation. These events develop the public sector's own functions and engage the staff to take action. The staff gets a chance to learn by doing and test service design methods as well as take part in pilot projects. One informant (A3) explains how these sessions have led to interesting opportunities.

> *'For example within the social and health sector, we've been training service design methods to hundreds of people. The training take four to five days. They've included mainly middle management, designers and people, who e.g. work in supportive functions. After the training, we've done pilot projects together with personnel, where we've engaged them. Those are the kinds of projects that you can't really see from the outside, and they kind of are development of their own operations. There are a lot of inner processes, that we've re-designed.'* (A3)

While the amount of projects has increased, also the scope of these projects is much more diverse than before. One of the informants (A7) describes how much of her present work includes organisational improvement and building of capacity, fortifying the service design knowledge in the public sector.

> *'We have been working a lot with the social service and welfare organisations. But also quite a bit with municipalities and local government on the municipality level, where they want to learn service design. So that type of work has been more directed towards capacity building, rather than addressing or doing certain project.'* (A7)

Also one of the informants (A10), who has been working in different government-related projects, described how capacity building has seen an increase in demand in the public sector:

> *'I mostly work in social and healthcare sectors, and it is all largely what I would call participatory design, with organisations or teams involved in the design of the service. It might be service, strategy, policy or whatever. I used to do more project work before, but now I do much more team capability building around them to do my kind of work, and helping teams in engaging qualitative research, and helping teams run co-design workshops with communities and stakeholders.'* (A10)

4.2.4 The Role of the Service Designer

In the following, the informants describe, how they experience their role in the public sector service design projects. The informants explained that the most important skill for the service designer was mentioned to be ensuring that there is a systematic service design process, committing and bringing together different stakeholders in order to make sure the project is a success. The human-centric methods of service design can be very disruptive for the public sector that's full of different interests, power and politics. The informants described that service designers have the skills, awareness, methods and know-how that's needed to navigate this sometimes very conservative field.

One informant (A4), who works in a service designer consultancy in Finland, emphasised the expert status of the service designer.

> *'The service designer has to be sharper than everyone else. Creating a cohesive and logical process is important, and it's important to communicate it clearly to the different sides of the project.'* (A4)

Also one of the informants (A9) emphasised that service designers are in a strong expert position in the public sector service design projects, by using customer insights as evidence by which to steer the project and organisation towards its key objectives.

> *'When we go and talk to clients and users, we learn a lot. And we accumulate a wealth of understanding of the customer experience. And sure, we can report all that we want, but all that information and knowledge will not be transferred to the organisation.'* (A9)

One informant (A3) explained that the expert role of the service designer strengthens gradually during the process, because of the customer insight that's based on the accumulated customer knowledge.

> *'Often there comes a point, where our position changes and we're not just experts in service design, but experts in the very customer experience they're offering. This gives us a head start to be able to tell the client, that this may be an expensive investment, but this would make your customer experience a lot better. Top management then has to make a choice, that if customer experience has been placed on a strategically high priority, then is it wise to ignore this expert's opinion?'* (A3)

The same informant (A3) described a model of three levels where the effective service design affects, as described in **FIGURE 11**. In the model described here, 'service level' is the level where the service is delivered to or experienced by the customer; 'system level' is the level where the resources are created, deployed and managed, with which the service is being provided; and finally, on the top level there is the 'strategy level', that guides the strategic direction of the organisation and gives the mandate to act.

> *'A sort of feedback loop is required, where the understanding is coming from the bottom up, all the way to the top, strategic level. It's important, that the strategic level understands what the customers think, how they feel, and then get the authorisation and the resources to trickle the process back down. The top level says to the mid-level, that the customer experience is important and allocates certain amount of resources. The mid-level says that we have now resources to make the customer experience better, and then the lowest level can utilise that understanding.'* (A3)

FIGURE 11 One of the informants (A3) described three different levels, where the service designer moves and influences the public sector service design project: service level, system level and strategic level.

The same informant (A3) explained that the above mentioned lowest level, where the service is solidified to a customer encounter, has traditionally been the strongest area of the designers.

> *'In the service level, we move within empathic design and participatory design, with the goal of getting both the customers and the staff to design together. But the top level is a big challenge and we really haven't been able to influence that. Now, however, they've started to listen to us, and we've been invited there more often. So we have a chance to influence the strategic level, but we ourselves have to learn this and to be sharper in what we're doing. Everything works in a slightly different dynamic, and there are a lot of tools of influence to make this work.'* (A3)

One of the informants (A6) also explained that in order to implement innovative and complex changes successfully, a high standard of conceptual thinking and strategic understanding is required from the service designers. This is very evident especially in those organisations where the orders come from the top and trickle down to the employees below.

> *'We have to learn to talk better and function more fluently within the strategic level. If we really want to make big changes and reassure, we have to understand that field and be able to adapt. In the same way, we've learnt to understand a lot of other operators and applied the human centric methods of service design. It works the same way. If the higher levels decide on the big resources, where there's big risk and it doesn't work, there's a lot of money to be lost. The stakes are obscenely high. Then again, if you botch up some service procedure, it's nothing compared to the big picture. Typically they seek certainty and fear failure a lot. But at the end, it's not possible to predict the future. You can only do some sort of analysis and draw some directions of where you should go to.'* (A6)

Many of the informants described service designers as the new change facilitators, working towards enacting change for the better. While change can be difficult, one one the informants (A4) suggested that the service designer has the capabilities to help the project onwards, towards the goal, help ideate alternatives, make conclusions and solve problems.

> *'The role of the service designer is to be an agitator and a facilitator, that directs people to another way of doing things. Steering them to do things differently and question the status quo, which can't be done, if there's no mandate from the management, that 'it's allowed to do and try'. It's a terribly critical factor – if a mandate doesn't exist, nothing will happen. No-one is going to even try. Then you sit with your arms crossed and say they aren't up to this.'* (A4)

Most of the informants explained that working with the public sector always requires the management's mandate, especially when working beyond organisational lines and silos. One informant (A5) highlighted the importance of being a facilitator and an agitator.

> *'Often the most essential part of development in the public sector, is about the operations between different functions. This means tearing down silos, to which you need at least two managers, who understand that the project needs to be done in co-operation. That they understand, that project is authorised, and that there is a licence to try things and cross governmental lines.'* (A5)

4.2.5 The Significance of Co-operation

In the following, the informants tell about the different forms of co-operation within the public sector, as well as the importance of engagement for successful co-operation. Service design covers many areas and disciplines, and a greater collaboration between sectors and different stakeholders is needed to address cross-cutting issues. Different people have a lot of bring to each other, and service design works best when people collaborate and make informed decisions together.

Most of the informants explained that the co-operation between different organisations and functions is important for engaging the different parties and for the project to succeed. One informant (A6) how a successful co-operation requires the client organisation to adopt new ways of working and develop the right mindset.

> *'Engagement is important in everything – bringing people together. It's maybe even more about developing the mindset. The fact, that an organisation that has a crappy service, would start providing good service, would require a change in the organisation. It requires a change in the organisational culture, and it requires change in the ways things are done.'* (A6)

One informant (A1) explains, that a shared vision is needed for achieving the goals the main outputs of the project, as well as to proceed and develop structures for co-operation.

> *'Organisations usually understand the value of co-operation themselves, and our role is to clear the way and to specify that co-operation. Co-operation is not an intrinsic value, but thinking over together, why and how we should be doing co-operation. So there's no superficial co-operation and we just sit together. It's important, that everyone sees the meaning and benefit, that's most often brought through customer insight.'* (A1)

The same informant (A1) continued, and explained that customer insights can help the organisation to understand who current and prospective customers are, in order to reduce overlapping policies and eliminate inconsistencies in the service delivery.

> *'Organisations should be made to understand that the customer does not e.g. have to visit several agencies, but there could be a new place built, where all the things could be handled. The social and healthcare sector is a good example in that even though everything is put together, it's still all a lightyear away from one another. Usually, organisations don't understand each other, don't speak the same language and don't see the same customer.'* (A1)

One of the informants (A5) continued along the same lines, and told that the traditional organisational structures of the public sector can bring their own kind of challenges to the co-operation.

> *'In the majority of public sector projects have the problem, that the organisations can be even physically separate. Bringing them together is often very challenging. Many people can be doing the same things without any knowledge of each other and there's a lot of overlap. Also pointless investigations, like situation reports, are done enormously. Everything is turned into an initial report.'* (A5)

One informant (A3) also told, that especially in situations where the need for change exists, engagement adds to positive attitudes toward change and commits employees to the process better.

> *'In one social and health sector project we went on location to get to know and talk to the staff. We talked a lot with the customers, too, and documented everything with a camera. At the location we had already a lot of ideas, but they were mostly hunches, that this might work. With this approach, we wanted to confirm that the whole staff was on board. That they will change their operating models, and we can get the management and the staff behind the project.'* (A3)

The same informant (A3) explained that, design games and gamification can help engage the staff to the design process in a whole new way. According to Vaajakallio (2012) and Brandt (2006), design games are not competitive but collaborative about stating participation, and they can be used for different purposes with various means and within different design fields.[182]

182 Vaajakallio 2012; Brandt 2006

> *'In one other project, we tried out a new approach to engage and do design. Gamified design, where we can engage more staff into the conversation and collect the views together. A bit of educating as well on the strategies of engagement. We did strategy cards and method cards so that they could watch what we do, what is a strategy and what we should still be doing and collect feedback after that.'* (A3)

The same informant (A3) continued, and explained that the design games are good in principle but also have their own challenges, being a new way of working for public servants.

> *'We've had positive feedback, that it's seen as a good way to work. But of course, because it's a new way of working, there are always people, that are more enthusiastic about it and those, who are less receptive. If it's the social and health sector, in my experience people are more open to new procedures. There are people, who are more open to new human encounters and contacts. Then the doctors are the hardest and most conservative group. For them, engagement isn't necessarily in the blood as much as the expertise is their identity.'* (A3)

One of the informants (A5) explained that a shift in mindset is the first step towards change.

> *'Right in the beginning of the project, we're making this kind of exercise with the client, where we look through their service together, from the customer's point of view. This is often the kind of thing that's never been done in many organisations. Really often the customer has, of course, not been looked at from within.'* (A5)

At best, a successful co-operation within the public sector can, however, give also new enthusiasm and positive vibe, that shows in increasing orders. One informant (A8) describes it as the snowball effect.

> *'Those public sector organisations that we work with, have accomplished so much, and have made a great change better for the public, and better for the employees. The organisation and the management likes working with us somehow, which means that other public sector organisations are looking at those aspects as well.'* (A8)

4.3 Defining the Problem and the Solution

In the following, the informants tell their views on the initial stages of the public sector service design process, where the work is guided typically by the defining of the problem and '*framing*' it right. The majority of informants described how it is crucial to formulate and frame the problem rather than trying to articulate the solution. According to Vaajakallio and Mattelmäki (2013), seeking the right issue to solve or the right question to ask, ought to dependably be done in coordinated effort with the key people who are developing public services to ensure long-term effects.[183] One informant (A2) explains that right in the beginning, the problem should be defined together with the client.

183 Vaajakallio & Mattelmäki 2013

> *'Problem definition is almost always done together with the client organisation's project team, who has power to make decisions and a mandate to choose the right problem to tackle it together.'* (A2)

One of the informants (A5) explains that there are no shortcuts to finding and defining the right problem. Through customer insights and qualitative research, underlying factors and drivers can often be uncovered in the initial stages of the process.

> *'There's no process to finding the right problem. It takes conversation. It's no use getting caught to the symptoms, but the diseases. The service designer's role in this phase is to bring understanding to the table and look at things together. At first, you need to understand, what it is that you're solving, rip open the process, find the key areas that we begin to investigate. Prioritise, which one of these is the problem we should tackle.'* (A5)

Service design as an approach can help to address complex social challenges that require systemic solutions. The informants explained that framing the right problem will not only help to better understand the underlying causes, but also to identify opportunities and turn these insights into relevant solutions and leverage points – services, tools, actions, and actors – that can transform the whole system from within. Thus, the service design process can be a useful framework for identifying those leverage points.

Looking at and understanding systems is not a simple matter, but it is the key to figuring out the underlying root causes of the problem. One informant (A1) explained, that finding the root causes of the problem through customer insights is essential to defining the problem.

> *'In the beginning, we focus on defining the scope of the problem and framing it. We'll strive to understand the real problem in the background and why it needs to be solved – and we'll also find and bring forth the root causes of the problem. During the customer insight phase, we're aiming to collect information, that helps to formulate a clear understanding of the scope of the problem. Understanding is built from user needs and motivations, through different methods. This knowledge can then be further refined with different user profiles and blueprints. The aim is to create a customer centric procedure, with which the service being built can be defined and produced.'* (A1)

In many cases the problem needs to be reduced to smaller, manageable problems. Once there is a deep and rich understanding of the system, the problem may look different from the systemic as compared to the individual point of view. One of the informants (A7) explained, that the problem can seem to be completely different from the points of view of the target group and the service provider, and sometimes the buying organisation has to be willing to change their own behaviour.

> *'We've been asked, for example, to build a completely new signing up service. After a few interviews we realise, that it's actually the internal governance of the organisation that is not working, and the solution to the problem requires changing the governance! It doesn't matter if you aim for the best service ever, if you're holding back your own employees.'* (A7)

One informant (A8) explains, that the focus of the project may change significantly along the way in a service design project, if the current focus is on the wrong problem.

> *'The scope of the request of the project is sometimes that they have internal needs, but when we come through the project after a while, we realise that there are other issues than they asked from the beginning. Because we are open, and actually want to achieve change, we have a dialog with the client to change the project. Some clients and customers are more open to this, others less open. But then it just takes time. You might just do the project that they are asking for, but in the long run they usually come back and try to make the difference for the future. Making a new request maybe, if they get the money from the different organisation or whatever.'* (A8)

One of the informants (A9) outlined that guessing or presuming is dangerous when defining the problem. Presuming is not based on fact or knowledge, but rather on the supposition of the way things are, of how people think. Presumptions are used when there is insufficient evidence or lack of knowledge from which a conclusion can be drawn.

> *'Presuming is the worst. The end result can be a terribly different solution compared to what was thought in the beginning. The presumed problem can turn out to be minuscule or completely wrong. Before moving into solutions, it is important to recognise user needs and understand the problem from different points of view, and also understand that the problem itself can have multiple, completely different solutions.'* (A9)

4.3.1 Qualitative Evidence

In the following, the informants describe, how they generate evidence, or customer insights through qualitative methods in the public sector service design projects.

There has been many attempts to define the concept of qualitative evidence. Scholars like Munro (2007) and Noyes (2007) argue that qualitative evidence is '*any study that is utilised both qualitative data collection and qualitative analysis methods*'.[184] According to Briggs (2007), qualitative evidence can be characterised as studies in which qualitative strategies were utilised to portray individuals' encounters and experiences.[185] Marston (2006) has even more broader description, defining qualitative evidence as '*any study reporting empirical, non-numerical data*'.[186] In the context of service design, all of these definitions are applicable.

184 Munro et al 2007; Noyes & Popay 2007

185 Briggs 2007

186 Marston et al 2006

Most of the informants explained, that the solving of the problem is typically started by engaging customers and end users to the process, as well as clients and other stakeholders that the problem concerns one way or the other. This phase involves interviews, field observations, experiments, that happen typically near the actual operational environments. The informants explained, that the qualitative methods such as casual theme interviews or more structured thematic interviews are typically used to generate customer insights. In these interviews, the topics covered are planned beforehand, but the order and form can vary.

One of the Informants (A9) explained that in-depth information such as customer's needs and wants can be only be collected by using qualitative methods.

> *'The qualitative component is an important part of the process. The more the process is exposed to the users, the better the end result is. The use of empathic methods, offers a possibility to understand the significance of different perspectives. With these methods, one can empathise and set yourself on the same level with the customers, and find and describe the human truth of the subject.'* (A9)

Qualitative research methods like interviews are essentially always based on human-centric principles and a strong customer engagement. Bowling (2002) argues that qualitative research is exploratory and descriptive in nature, following an inductive process in regard to how information is gathered, analysed and interpreted.[187] One of the informants (A4) explained that the purpose of these techniques is often to validate assumptions quicker, spot the obstacles that need to be removed along the different stages that customers go through when interacting with the service, and also finding opportunities to create value.

187 Bowling 2002

> *'Through the interviews, it's possible to find out, where the customer feels pain the most. And how to create relief. Through the interviews, you can get additional feedback to your own assumptions on obstacles of service use. Obstacles are the problems that need to be eliminated or designed away.'* (A4)

Almost all informants explained that interviews are an essential part of the service design process, allowing you to talk to people to express themselves and giving you the benefit of adapting your questioning based on the responses that you receive. One informant (A6) described that usually there's a 'critical mass' in terms of the number of of interviews and interviewees.

> *'After 10 to 20 interviews, the answers start easily repeating themselves. Ten interviewees is easy to control and recruit, and usually the interviewees are quick and easy to pull from e.g. the existing clientele.'* (A6)

In service design, qualitative evidence is often collected and generated through qualitative methods or experiments that are low-risk in nature and is based on the experiences of the target group. Qualitative methods such as interviews help to understand how people feel, do and think, or why people do so. Two of the informants (A5, A6) explained, that after a small round of interviews, patterns start to emerge that help in the formulation of hypotheses to understand if a proposed solution creates value.

> *'In the first stages several qualitative methods are used, whose purpose is to understand the problem and its complexity. For example, if you take sampling of 10 to 20 people and you ask them, that if this kind of service would exist, would you be interested? You ask them, do you feel this is relevant to you. Is this good content-wise, is it communicated correctly. After 5 to 6 interviews you can usually see some type of a trend. It's always very inspiring to discover the existence of the trend.'* (A5)

Torres (2017) argues that direct questions like 'would you use this service' are not good at speculating about customer's future behaviour and the answers to these questions are not reliable. According to Torres, asking more specific questions and getting people to take action rather than asking them what they would use the service. One informant (A6) explains that this is important especially when designing services for non-digital savvy people.

> *'If you talk about people that don't use e-mail of digital services, the interviews and hypotheses are a way to investigate things like, what could be the most value adding view to have people use the service. Trying to squeeze everything out of people. Then you also observe and guess the rest. The information has to always come from the field. It's important to seek answers – not settle for assumptions.'* (A6)

While the nature of qualitative evidence is not scientifically rigorous per se, the qualitative methods provide rich, meaningful and actionable evidence, hints and cues about the customers needs and concerns. One informant (A1) explained, that a holistic, qualitative understanding of the phenomena functions as a form of evidence, to which you can actually base your decisions on.

> *'Customer insights work as a kind of evidence in the problem defining, which can be helpful when determining, why this is a real problem. If the real problem doesn't seem to be found, you go out and test, if it's a real problem. You go back to the customers and ask, whether this feels like the kind of problem you bump into when using this service.'* (A1)

One informant (A7) described that quotes derived from customer interviews are a strong carrier for customer insights, and represents the needs and aspirations of a wider population.

> *'There's has to be a balance. Because otherwise if you have – not an enemy, but a sceptic person amongst the client's organisation, they would say that 'It's only one person who says this — we can't change this.'. All the time we always need to have a balance. Yes, here we are telling one story, or one quote of a specific person, but we are choosing that because it represents what many people feel. The way this specific person expressed it, really expresses the concerns of the larger audience that we have revealed here. So that's always balance when using quotes.'* (A7)

4.3.2 Generating Evidence with Hypotheses

In the following, the informants describe how they test hypotheses using qualitative methods, and how they collect and generate evidence through validating hypotheses.

According to Rogers (1966), hypotheses are good hunches, distinct cautious speculations, applied in formulating theory or planning an experiment, projected to be given a straight experimental test when possible.[188] Philosopher of science Karl Popper (1959) has argued that falsification of hypotheses is possible, and verification is impossible. According to Popper's concept of falsifiability, a scientific hypothesis must be testable, yet there is a considerably stronger necessity that a testable hypothesis must meet before it can truly be viewed as scientific. Popper argues that it is not difficult to accumulate evidence for just about any idea, but a hypothesis is useless unless it is '*risky*' – it must make predictions that could negate or contradict it.[189] Thus, a valid hypothesis is falsifiable, that is, it makes a predictions or statements that can be tested helps to answer the question: '*what are we hoping to learn?*'.[190] One of the informants (A1) explained, that these hypotheses, once well tested and verified by experimental data, become more concrete solutions that can be then taken to a more concrete level.

188 Rogers 1966

189 Popper 1959

190 Rusonis 2015

> *'Many organisations start with thinking, they know what the problem is. To understand, there has to be a hypothesis about the problem, to understand what the solutions could be. When we get evidence about the functionality of the solution, that this thing works, we usually start scaling the idea. Make it more concrete, scale it and take on some more recourses.'* (A1)

Service design is all about co-designing and co-developing ideas together with the customers and different stakeholders. Most of the informants explained how they utilised the hypotheses directed at the target group, as a way of testing, evaluating and validating potential solutions. The informants said that hypotheses accompanied by visualisations also help to make ideas concrete, as well as help them figure out the scope of the problem and to map out the needs, wishes and wants of the users. One of the informants (A1) explained that in the beginning of the project, the most vital thing is to concentrate on the right customer, the right problem and finding the right solution.

> *'First, you have to understand, who the customer really is, which is usually not a simple question in the public sector. What are the customer's needs and why are they coming here. Then, when you understand the need, there lies the problem as well. You understand, that this is the problem you're solving. And you understand, how the problem is being solved now, and how it should be solved better. In a way the object of development, that's going to be tested through hypotheses.'* (A1)

Hypotheses have their roots in observations about the world along some assumptions, by which a certain outcome is predicted. According to Rusonis (2015), a well crafted hypothesis articulates clearly what is being changed, *what* is believed that the outcome will be, and *why* is it believed the outcome will be, and *why* that is the case.[191] O'Reilly (2016) argues that in hypothesis-driven development, learning happens through investigation, acquisition of information, and correction of assumptions. The experimental approach to information discovery includes a series of steps: observation, hypothesis formulation, design of experiment to test hypothesis, determining indicators of success, conduction the experiment, evaluation of results obtained from experiment, accepting or rejecting hypothesis based on indicators, and finally, formulation of new hypothesis.[192]

191 Rusonis 2015

192 O'Reilly 2015

One informant (A9) explained, that using an experiment and whether the hypothesis is confirmed by the experiment or not, learning happens through iteration, which may lead to other hypotheses in turn.

> *'Hypotheses are informative guesses or concept ideas to be tested, to tell what the service being built, or parts of it, could be like. We create hypotheses, and show them to the customers right away. Ask them things like 'What if there was a service like this…', 'Should there be a service like this…', or 'Does this have any idea…'. Hypotheses are taken away or combined during the process. They can also be created on the fly. They are also iterated all the time, constantly.'* (A9)

Hypothesis-driven approach to product and service development was introduced to the mainstream audience by Eric Ries (2011) in his book '*The Lean Startup*'. Ries outlines that hypothesis-driven development happens through a series of tests that will be validated in production. The outcome of those tests and the validation of the hypothesis will provide the essential feedback to continuously improving the final solution.[193] According to Barry O'Reilly (2016), who has additionally been building up a hypothesis-driven development framework for software development, it is vital to ensure that the hypothesis is falsifiable, vividly, it can be clearly checked or invalidated through a decisive experiment. According to O'Reilly's framework, it is imperative to view experiments as processes to test hypotheses about the system. Learning is crucial for making the results from these hypotheses useful.[194] One of the informants (A9) explained, that it's imperative to understand, what the value of the hypothesis is from the perspective of the customer.

193 Ries 2011

194 O'Reilly 2016

> *'What's important about hypotheses, is that they are not meant to be realised as is. But to understand, what is the value of a hypothesis. Why will the customer or employee value it? The final concept will be created through values. It might include elements of the hypothesis, but the most important thing is to*

> *understand, how to create value. Hypotheses also have to have metrics with them, because that puts a value to the conversations. Numbers steer the work in every phase and every discussion.'* (A9)

One informant (A4) explained, that before moving to the actual design part of the service concept, early ideas and potential solutions are a good way to make things concrete.

> *'You can't get very deep into the user's everyday life or habits during fast interviews, but through discussion, you quickly get an initial picture of the possible uses and bottlenecks of the service being built. The earlier you can make the pieces of the service more concrete, the better the potential users of the service get the idea.'* (A4)

Ohala (1986) outlines a process, how hypotheses and evidence interact with each other. First, hypotheses are prepared then evidence is brought forward for and against them. Some of these hypotheses are acknowledged and gain a following, after which they get succeeded by new hypotheses that have better evidence.[195] According to Ohala, what evidence accomplishes in this procedure, is to guide us to pick between competing hypotheses. Rather than approving or demonstrating one hypothesis, it discredits competing hypotheses. Two informants (A1, A9) explained, that in their simplest form, hypotheses can be rough visualisations of the service being built or parts of the solution:

> *'Hypotheses are tested by experimentation. In the problem defining phase this means basically, that you interview people and show them thoughts. You can make like 20 hypotheses and make them more concrete through rough drawings. After that, you go and meet customers or employees. Tell them the operational context and show them the hypotheses. In this stage, we don't want to know, if they're good or bad, but why they are good or bad. Solutions will start to be eliminated only after the problem has been defined.'* (A1)

> *'The first hypotheses have to be quite rough visually. They are thoughts, that are brought forward. The purpose is not to present solutions, so that the interviewee doesn't lock themselves into their thoughts. It's easier to discuss abstract thoughts. If you take a finalised layout to the customer too early, the whole conversation is about the layout. It's a challenging skill to do visualisations, that don't lead the conversation to appearance or usability.'* (A9)

Hypotheses can be considered as informative guesses that provide a way of framing question such that they yield specific predictions. Two informants (A4, A8) explained, that there are also different approaches to gather understanding and gain new perspectives on the matter, such as more conversational approaches, instead of testing hypotheses per se.

> *'There can be found whole new angles from interviews instead of just aiming to validate an existing hypothesis. That's why it's important to learn to let go of your assumptions and prejudices.'* (A4)

> *'We need something – a foundation of subjects – that we talk about. Which is something that the client is interested in. They know because they have quantitative data, figures, numbers and different things – and we want to talk about this area. But we don't do clinical research per se, in the form of 'we have a hypothesis and we want to see if this hypothesis is correct or not'. Because then we are painting ourselves to the corner. In a way, you just have to trust the process more - we just go out there and see what happens.'* (A8)

One informant (A4) explained that validating the right solutions through hypotheses and experimentation is crucial especially when developing and designing digital products and services.

> *'The wrong assumption, or uncertainty of whether this is the right solution, is the most fundamental mistake one can make when developing a digital service. Then if you are sure of it, the development of a digital service continues on as long as you like, going into developing usability and how to make it smoother, prettier, amicable or educational. It's an art of its own, that continues from here until the end of times, because technologies develop and equipment develops.'* (A4)

According to O'Reilly (2016), hypothesis-driven development in software development comprises three key elements: functionality of the experimental process on the hypothesis, experimental results, and last but not the least, indicators of success. To make sense of the statistical data, effective monitoring and evaluation tools become a necessity. Another factor to keep in mind is that the assumptions for the hypothesis need to be clearly stated to create a sort of feedback loop and criticism. A/B testing, customer surveys, beta testing are all examples of hypothesis driven development mechanisms that generate different forms of evidence. Combined with continuous delivery, such feedback expedites learning and experimentation for development. Hypothesis-driven development is hence an opportunity to guess the nature of the problem before embarking on a solution.[196]

4.4 Developing Through Experimentation

In the following, the informants describe how they utilise experiments and experimental development in the public sector service design projects. The informants told, that experiments are about e.g. testing a new idea, validating a solution, bringing a service to market or scaling a service. Each one of these requires a different kind of testing with different customers and stakeholders.

According to Buijs (2007) and Mumford et al (2002), the innovation process is the presentation of novel ideas comprising of a progression of stages for idea processing, including evaluation and, subsequently, implementation.[197] Often, solutions begin to emerge on a trial and error basis of experimentation. Hence experimentation is an important part of innovation, and becomes all the more relevant when dealing with complex and unclear issues.[198]

197 Buijs 2007; Mumford et al 2002

198 Argote & Ingram 2000; Andriopolous & Lowe 2000

Most of the informants explained, that in order to foster service innovation in the public sector, experimentation and experimental techniques are used to manage change. The informants mentioned different aspects that should be considered when experimenting: fidelity (or similarity to real conditions), cost of experiment, iteration time, capacity (or realistic estimation of number of experiments possible for fidelity), sequence of steps, and type of experiment. One informant (A2) described how experimentation is utilised throughout the service design process.

> *'Experiments are, in a way, part of the design process for us. You could say, that it starts, when the client organisation is hanging around in the concepting phase. We make a lot of co-designing and customer testing along the way, so it becomes more like doing. It can mean for example, that you try new procedures, and then if it works, you expand on it some, and so on. It's pretty much like including the gang, sort of. And on the other hand, it's defining what at the end of the day, is succeeding and what isn't.'* (A2)

Experiment-driven models for innovation differ marginally from conventional methods, which lack iterations, and posit pilot test launch for projects. The iterative process in the experiment driven models, however, provides a feedback loop for analysis and improvement. Using prototypes, which are critical for innovation, the project's cost can be reduced, and the chances of success are raised.[199] One of the informants (A5) described experimentation as an iterative process, where the service finds its final form through the process:

199 Thomke & Reinertsen 1998

> *'The experiments are not about testing the functionality of the final service, rather the service will mould and adapt through the experiments towards its final form. Only experimenting and testing you can find out, which solutions work and which do not work.'* (A5)

The same informant (A5) continued and explained, that experimentation makes it possible to simulate the desired experience, make solutions more concrete and show their functionality:

> *'The main point of the experiments is to make things that are hard to understand, simple and easy to grasp. Show, through practice, what is good in the idea and what requires development.'* (A5)

Experimentation is a fundamental method to learn from uncertain, distant outcomes, where the amount and value of learning determines the success of experimentation.[200] Improvisation and creative application of analytical concepts yields new ideas, and the key factor is determining what works and what does not. One of the informants (A4) explained, that experiments are an useful way to try and create lasting value through learning:

200 Thomke 2003

> *'Experiments develop the beginnings of a solution further, by using feedback. We're trying to create lasting solutions through experiments, that create value in a lasting way. You can also commit the different parties needed in the process very cost efficiently. This can be necessary for the project's success.'* (A4)

In experimentation, targets should be set, and by experimenting with the least possible resources and playing with system variables, tries to learn from the target. According to Torres (2017), there are also situations where experimentation should focus also on simulating the desired experience. According to Tuulenmäki and Välikangas (2001), when implementing an experiment driven approach, three different idea types need to be kept in mind: opportunity idea (the idea imagined for problem solving), experimentation idea (testing critical assumptions related), and execution idea (the information from experimentation).[201] One informant (A6) explained that dividing the big picture into smaller pieces is important for controlling the complexity:

201 Tuulenmäki & Välikangas 2011

> *'The initial idea of the solution could be quite big, but the experiments have to be small. The big picture has to be able to be scaled down and understand, how this could be experimented or start off even if there would be a big service entity running in the background. Through experiments we have to be able to make sure, that we're solving the right problem before the idea is taken further on a concept level.'* (A6)

4.4.1 Learning from Experiments

In the following, the interviewed service designers describe, how they plan, measure and evaluate experiments. Almost all of the informants utilised experiments to make the ideas derived from customer insight more concrete and to test them with end users, as a way to support reaching the greater goal.

As an instrument of experimentation, fast prototyping is a fundamental approach to gain from the iterative procedure of experimentation. The iterative trial and error yield new information, which, combined with earlier knowledge of the subject, provide insight and competitive advantage.[202] The informants described that different stages are involved in the experimentation process: *hypothesis, planning the experiment to test the hypothesis, executing the experiment,* and *analysing the results*. One informant (A6) emphasised the importance of well planned experiments, and that the experiments should be a way to also learn.

202 Thomke et al 1998

> *'Experiments have to be planned and executed well and in an organised manner. This way we can be sure, that the experiments and service prototypes lead to usable information and they can be learnt from. Learning is always at the centre of the experiments.'* (A6)

Two informants (A4, A7) explained, that failing is a good thing, because then the experimented solution doesn't take you to the expected end result.

> *'The same thing can be tried out in a number of ways. Failings are always allowed, because experiments are a tool for learning. The longer in the process we are, the more important it is to focus on trying out the right things.'* (A4)

> *'The first experiment as such doesn't necessarily lead to anything, it's just tried out through developing. Often times the case is, that when an experiment is set up for the first time, after that the real problems of the service being thought of and the production of it, start to reveal themselves'* (A7)

One of the informants (A1) explained, that rapid experiments can be used to validate functioning solutions and drop out bad, defunct ones.

> *'If the hypothesis or the presumed solution doesn't work, it's worth abandoning it quickly and move on to the next solutions. Tested and proven solutions are worth developing further.'* (A1)

The same informant (A1) continued, and explained, that the biggest reason to do experiments, is overcoming pitfalls.

> *'There are always new challenges coming your way when experimenting. But the idea of these experiments is to tackle these challenges. You should be able to adapt the experiments to the point of being pretty sure, this should work. This is especially important for digital services.'* (A1)

One informant (A9) explained, that failing makes learning possible.

> *'Failings are part of experimental development. Failing is a result, after which you should of course think about what failed and why. On how both the idea and the procedure could be developed through the teachings that come from the failing.'* (A9)

Two informants (A2, A4) explained, that when it comes to rapid experimenting, quality of experiments is more important than quantity.

> *'When you talk about qualitative doing, the amounts are pretty irrelevant. It's more about, what is the number of variants*

> *that brings you something new to learn, and where is the point where you've reached your saturation point. If you're looking for numbers, and you have to include lots of people, it should be done in a different way than experimenting.'* (A2)

> *'If you think about experiments in terms of testing, then e.g. 5 to 8 people in a usability test is a good, basic number, after which more people won't bring much new info in relation to the effort it would take, and you should change the thing that's being tested. Usually it goes over with a lot less, e.g. a co-designing sessions.'* (A4)

One informant (A1) explained that impact at the micro level must be examined more closely, and to measure those, a variety of different indicators can be used.

> *'When measuring experiments, more subjective metrics are used, like do you feel, that you'll save time, do you feel, that it's easier, what e.g. staff thinks among themselves. What has to be accounted for, one way or another is, do the customers feel the service is smoother. For that, there are different tools that, in essence, are questions with which you can evaluate, if the service has gone to a better direction.'* (A1)

One of the informants (A2) also explained, that metrics have to be able to assess continuously during the process.

> *'If the pre-set metrics are shown to be in some way inadequate or otherwise impractical, you might have to retort to another kind of evaluation. Were the metrics set for the experiments right? How well can the results of the experiments be generalised? What kind of results was actually gotten from the experiments?'* (A2)

4.4.2 Validation of Service Prototypes

In the following, the informants describe how they build and evaluate service prototypes to accommodate the evidence gathered during user tests and experiments at various stages of the service design process.

According to Blomkvist et al (2012), service prototyping is an imperative piece of service design practice that permits designers to learn fast, gather feedback and communicate about design ideas with stakeholders. Blomkvist et al present two viewpoints for the learning purpose of service prototypes, *exploring* and *evaluating*. Exploring refers to generating and gathering customer insights, while evaluating refers to testing of the developed solutions, receiving feedback and spotting points of failure.[203] Most of the informants identified these two approaches, and described how different service prototypes are designed and developed during the process to serve these learning purposes. One of the informants (A5) described how prototypes are used for testing purposes as learning and planning aids, as well as visual aids for communicating complex findings and ideas to different stakeholders.

203 Blomkvist et al 2012

> *'There's no shortcut to good service. We'd rather do prototypes than make presentations. Learning and findings are more important than the process itself. Is there a reason to look at the problem again? Has some prototype or validation brought more info that would suggest we need to revisit it?'* (A5)

The informants described how the prototypes developed in the early stages of the process were typically low fidelity in nature, and their primary mission is to articulate the vision and strategy using visualisations, storytelling and scenarios. The prototypes continue to develop during the process, and eventually start to come to life at a higher level of fidelity.

One informant (A8) explained, that the prototypes can be separated into two different workflows – strategic and operative.

> *'In the beginning of the project, the delivery, the prototypes and our concepts are more strategic. The further along we work with the client organisation, they become more accurate and detailed. Maybe they can be separated into two different workflows, as we have strategic areas that we have to work with. In the beginning it might be just concepts made internally that we show to customers to see how they react. But then we try it in practice. Or the customer tries it out without us sometimes as well, and then comes back to us to talk about his habits.'* (A8)

The concrete prototypes help understand also the other dimensions of the service being built, like the digital solutions. One informant (A1) explained, that developing validated prototypes into a minimum viable product (MVP) level makes early release and collection of real user feedback possible.

> *'We have more and more developers, that are capable to build a MVP type of prototype within a day, through which you can test its functionality in the real world. The prototype can be ugly and stiff, but it serves its purpose. If it starts to work, only then the more systematic building of a digital service, coding and such.'* (A1)

One of the informants (A5) described how empirically tested, measured and validated prototypes accelerate every phase of the process, yielding recommendations for further development.

> *'I'd say, that they become more sophisticated, the further the process advances. At first they can be just written scenarios of the different functions of the service. Then those scenarios can be visualised. After that, there can be interactive elements. And eventually, they can be true pilots, where the prototype*

> *itself resembles the finished service, and it's used by the client organisations personnel. Validation is performed all along the way to avoid pitfalls and seize the opportunities.'* (A5)

One of the informants (A9) explained, that physical prototypes help deepen the shared understanding, from the viewpoint of interaction and customer experience.

> *'The service processes happening in the customer interface, require the presence of personnel or the customers. Customer service is tightly connected to physical services, and it can be difficult to produce these kinds of services without physical spaces. For example, the services at the healthcare centres can build from several physical touch and contact points in the form of spaces, equipment or other concrete objects. It's easy to build a prototype from them, with which you can collect feedback as well as observe action and interaction.'* (A9)

Service prototypes are one way to learn to understand the potential users', like the employees' or customers' needs, preferences, and behaviours better. One of the informants (A3) explained, that physical prototypes can help assess the functionality of the solution in the customer interface, where an organisation meets its customers.

> *'Service delivery in the public sector requires often physical spaces. They almost always are about human interaction, and as a service, they come with some sorts of goals or needs, whose actualisation is important. From the customer experience perspective, with these kinds of services, it's important to understand e.g. service area use of space, functionality and practicality. This way the success of the solutions is easy to notice by testing a physical prototype.'* (A3)

With service prototypes, you can test and validate whether the developed solution works like they meant to be or are thought to work. The same informant (A3) explained that service prototypes are often experimental in nature, and let the stakeholders to try out and compare different solutions in action, helping to boost employee buy-in and engagement further.

> *'In one project, we built a physical service space from cardboard , where we had all our ideas laid out, made really roughly. Then we went around the space with the client organisation's staff, looked through all the ideas and marked what was good and what needs to be changed, what to take further. This was evidence in a certain sense, how we convince the client. Engaging and thinking really thoroughly the engagement. So that they can prove it themselves.'* (A3)

The same informant (A3) continued, and concluded, that the development of face to face customer service and communication is, however, challenging.

> *'The physical world is significantly harder to develop than the digital world. Pure face to face service is really difficult to develop, because it's so human centric. It can't be standardised in any way. Nothing ever happens the same way twice in interaction between people. There are things that can be rooted out or add to make it as easy as possible. Often times the thing that's added, is the digital dimension, on some level.'* (A3)

4.4.3 Experiments and Prototypes as Evidence

In the following, the informants are describing how you can learn from experiments and how to utilise the results further. The informants told, that the information gathered from the experiments can be used as evidence that the solution if functioning or not, or if the changes are useful and their impact before their broader implementation.

One of the informants (A2) explained that it's important to set clear goals for the experiments.

> *'First you have to determine the needs and the ideas, that can lead to the set goals. It's important to recognise, what is being tested and why. The better the experiment is confined, the easier it is to understand the results of the experiment and convince others of the functionality and usefulness of the solution.'* (A2)

The same informant (A2) continued, and explained, that clear metrics and goals guide the direction of the experiment.

> *'Experiments have to always have metrics with them. Clearly, before the experiment, set metrics are a good way to monitor the success of the experiment and based on that, do analysis and conclusions on its functionality. It the metrics prove to be successful, the solution can start to be developed further. If you can't make the solution work, realising it in a later stage is a lot more challenging, and more expensive.'* (A2)

One informant (A1) continued on the same subject, and emphasised, how there should always be metrics set.

> *'There are usually some kind of metrics and goals set for the first experiments. The experiment has to produce these kinds of results, that we can be satisfied and move on. There are always some metrics for an experiment, and with those metrics you have to believe, you can develop it.'* (A1)

> *'A failed experiment is an experiment, that has no goal and no metrics. Experimenting for the sake of experimenting, or starting just doing something. No-one can tell anything about was this a good thing or not. Sure, people often feel like the experiment had some sort of impact, but it doesn't necessarily steer the development to any goal-oriented direction. It's just nice dabbling, and the fact that 'we had fun' can be a kind of a result. That we had variation or something - persistence remains to be seen'* (A1)

One of the informants (A6) explained, that behind the different public sector experiments have typically a some sort of an assumption of how the experiment would impact the system at large:

> *'A successful experiment can have an impact for the whole organisation, even though the specific experiment happens on a micro level in a singular service function. Or if you consider broader strategic level experiments, they are based on the assumption of impact on a broader system or social level, even though the experiment itself would be strictly limited.'* (A6)

One informant (A6) explained that when talking about small experiments or prototypes, it can be impossible to manufacture strong evidence, or so called 'hard facts', or see the causal relationship between experiments and value creation. The informant described that impact has a completely different meaning within healthcare, compared to the context of service design:

> *'The biggest challenge is especially, when doing projects for healthcare or social and health sector. For them, impact means something completely different than for us. And understanding, that their impact means, that when you invest in the managing of people's overweight, the diabetes cases go down. How things impact on a larger scale. In everyday work, this means that you start to focus on every doctor or nurse intervenes if someone is overweight, or at least mentions or recommends some procedures. And the way this work shows is, that you can some day state, that cases of diabetes have gone down. In health sector that is impact, the kind of macro level change, and they're used to that kind of world.'* (A6)

The same informant (A6) explained that experiments that are developed inside a very particular set of variables and constraints that are not necessarily generalisable outside of the experiment.

> *'That's the challenge with evidence. It can be sometimes hard to extract and reuse in really straightforward ways. So, it's not like you can easily go and pick up the evidence. Sometimes there is no evidence. Sometimes there's totally new ways of doing stuff and there's no evidence to demonstrate what works. So, we use the closest most useful thing, we will patch some stuff together and we will just make sure that we are learning and building evidence as we go. So it is not always really straightforward thing.'* (A6)

Health is the space where strongest evidence occurs because it's the greatest risk, the most obvious risk. The idea of evidence exists in different places, but it is particularly strong in health. One informant (A10) explained that, especially in healthcare, it is more than dangerous for designers to use normal processes where the laws of accountability are very different.

> *'In the traditional design context you don't have to demonstrate your evidence. As long as it works, and 'works' is subjective, the clients are profiting from it, and people seem to engage in it. It doesn't even have to really matter if it doesn't work – 'Hey, it sort of works!' – people were not harmed. But in the health context if it doesn't work the way you thought it would, people will stop the project, because they can't have the same accountability.'* (A10)

The same informant (A10) continued, and explained that the evaluation framework has to be rigorous as the actual behaviour change model in healthcare setting.

> *'Some of the health-based projects – use the term 'evidence-informed'. And that's their way of saying, of course we need to look at the data that is available to us, but we're not tied to a really strict notion of 'evidence-based', because it is not helpful for the type of progress and systems change that we are going to do. We are not going to be 'evidence-based' and we want to develop evidence through lived experiences. So the term is intentionally 'evidence-informed' to give them more flexibility. I don't think that we should be doing any work in health and social sector that isn't at least evidence-informed!'* (A10)

> *'In one project we spent a lot of talking about what do we mean by evidence, and we demonstrated that lived experience is also evidence. When you say it's 'evidence-based', you should mean, in my opinion, evidence that comes from different sources. Different kinds of evidence. Lived experience is a form of evidence, and that's why I would start working with health professionals and projects, and I would work them through how lived experience is a form of evidence and how it becomes a form of evidence. But that's not what traditionally it would be meant by - evidence-based.'* (A10)

One informant (A5) explained, the design projects in the health and social sectors can be looked from two, quite different perspectives. When working with hospital environments, there is accountability to the public service and accountability to question, like 'How well are we serving the customers?'.

> *'When you are improving people's experiences in the waiting spaces, and helping them to get around the hospital, that is not an intervention, and we are not probably going to do any harm. We are going to do it worse or better. But we are not interfering with the health outcomes. The better the experience is, the better the outcomes will be, because there is less stress.'* (A5)

The same informant (A5) continued, and explained that when working with health interventions, or developing an actual behaviour change intervention, traditional clinical health model has to be integrated with the qualitative design model, and there is a need to look at both sides of accountabilities.

> *When you're doing health interventions, let's say that we are developing an application that we believe will help young people tackle obesity – it's like an intervention and we are directly interfering with the health outcomes. Then, you need to have as much evidence base as you can.'* (A5)

4.5
Validation and Impact Assessment

In the following, the informants tell their experiences about validation and impact assessment, what the successful public service design project is like, how success is defined and what is meant by impact within the context of service design.

Two informants (A7, A8) explained, that impact is the key driver in the public service design projects, that is set in the context of the impact on citizens, stakeholders and their relationship with the wider society.

> *'We want that the inhabitants of Sweden would have better everyday lives, and they really understand the democratic system and they feel that they can access it. And also that it's easy and accessible. And the employees as well – they are proud of their work, they have a nice work day, and they deliver services that they stand for to the inhabitants of Sweden.'* (A7)

> *'We really aim to achieve something for the clients, not just do the project. We listen to the customers and talk to the clients to change the scope to actually achieve something that is valuable to the customers as well as the client.'* (A8)

Service design involves both qualitative and quantitative methods. When things like emotional elements of the service are being evaluated, qualitative methods such as interviews and observation are relevant. Most of the informants described that numbers, or other indicators to measure success, are always needed to answer a question, and that is where quantitative methods come into play.

Well-defined key performance indicators (KPIs) can characterise the achievement or failure of the service. In the context of the public sector, indicators can measure the economic or social value added; economic indicators can be related to things like operational and efficiency measure, while social indicators relate to the context of social change. One informant (A2) explained, that different indicators should be used throughout the process to indicate where the solution is succeeding.

> *'Whether they are economic indicators or whatever – they have to be included in the project at some points. Private sector projects are easier, because they are always concerned with maximising expected euro returns, which are easier to follow. In the public sector, they talk about saving time, improving workflows and so on - which is of course also about saving euros.'* (A2)

When it comes to social indicators, some of the informants described how they sometimes face the difficulty in creating useful indicators in evaluating impact. This is particularly evident when dealing with complex or systemic problems, where understanding and analysing a diverse range of impacts can get very complicated. One informant (A5) described that in order to capture the emerging evidence of impact resulting from the intervention, you need to look at the right indicators that are simple and well-defined.

> *'It is crucial to find a number that defines success. However, in the public sector, this can be tricky. The most difficult part is to identify the right indicators, especially in the case of complex projects. Often in these cases, the data is too complex or incomplete in the sense that there may be missing data or errors in measurement. From my experience, indicators that are relatively simple and well-defined can be measured most effectively. Sometimes longitudinal studies, that keep going for a period, fill the need of approval of the intended impact. In the perfect circumstance, it can be inferred that there is a statistically and factually noteworthy difference regarding how the situation has developed in connection to the time before the intervention.'* (A5)

One of the informants (A2) explained, that success can be measured by comparing to the current situation or through statistical change.

> *'Comparing to the current situation is one metric for success. We have had an ongoing public sector project for a few months now, and we can't tell yet, if it's a success or not. But how we're going to evaluate the success, and how the client will also evaluate it is, whether it creates a statistical difference compared to the current situation, whatever the solution will be. At the moment the statistics show quite large differences, and it should start showing going in the other direction. The gap should start to get smaller.'* (A2)

Impact evaluation can be carried out also at the micro-level. One of the informants (A5) explained, that a successful service design project can be seen as improvement of customer satisfaction and more effective and efficient resource management.

> *'Of course, it depends a lot on the case. But we do think it's very important that the customer experience improves. That's our starting point. And with that, other advantages can be accomplished. That we can e.g. create processes, that has a bigger role for the client. That saves the organisation's resources. If we design a good experience for the social and health sector client, so that the customers understand, what's happening to them, the whole process becomes more efficient, when they don't e.g. have to call customer service all the time. That way a lot of recourses on the support side can be saved.'* (A5)

Measuring effectiveness is also about measuring how well the services meet customer needs and requirements. But also, how happy the customer was with the customer experience. As one informant (A3) explained, instead of focusing just on doing things faster or more effectively, the organisations need learn to to listen to customer feedback and use their input to adjust the customer experience.

> *'What service design process can bring to different service environments, is different metrics to guide service production - where the employees focus on at work. Public services measure effectivity a lot. But it's a completely different thing, whether to measure, how fast you do something, than for the focus to be moved to whether or not the customer was served. One critical thing in the development of a service is, that if an organisation is able to really listen what the customer thinks and takes it into the service output in a different way. What are the metrics and how is the staff guided to think whole new thoughts.'* (A3)

One result of the process can be to enrich the public organisation's existing indicators with qualitative data. The same informant (A3) continued and explained, that the functionality of a solution should be assessed constantly against different criteria, even if things become more efficient.

4.5.1 Implementing Organisational Change

In the following, the informants tell their experiences about implementing organisational change. Many service designers felt that implementing change in public sector organisations was difficult, but not impossible.

The informants portrayed that the move towards more customer-centric practices such as service design starts with an adjustment in the way of thinking that includes replacing the typical development and operating models with more empathic methodologies. Customer-driven innovations emphasise the needs of users and empower the participation of citizens in the design and development processes. It may be the case that, upon a strict interpretation of relevant law and regulations, end users of the service – citizens – are not considered at any given point of the public service delivery. As soon as the people who use public services are forgotten – living, thinking, feeling human beings – things go wrong. After all, public services are relationships between service providers and people. One of the informants (A3) described that different strategies are needed to engage the top management to recognise the importance of customer-centricity.

> *'During the process, we're moving more in the higher level in decision making. We're trying to always ideate all kinds of tools to guide that thinking into more customer-centric action and that they would focus more on our solution.'* (A3)

One informant (A4) also explained that engaging the staff and getting them to commit can be a key to implementing change.

> *'Some services are more people friendly, and change will not happen from top to bottom. You have to be able to influence people's behaviour and way of thinking in a positive way. The*

way a message is communicated to personnel and the language on how it's communicated is important. If the solution is not connected to the organisation, the change will probably not happen.' (A4)

One informant (A6) explained, that implementing change usually requires convincing the client organisation.

'For many people, service design process is a new thing, and questioning it is a typical change resistance strategy in all levels of public organisations. This can lead to underperforming in meetings, where things can't be discussed as deeply as the situation requires. In these situations, the client can be convinced by showing selected evidence, like authentic user quotes, pictures of service situations and other material. This way the client organisation gets to the root of the problem, understands the customer's needs better, as well as the value of the solution. This helps in building strong trust and dedication.' (A6)

The same informant (A6) continued, and explained also, that evidence can have a significant role in implementing change:

'The customer insight work done during the process functions as evidence for the functionality of the solution, and works as both a sales tool and as support for the actual decision making process. As far as successful project goes, it's important that all parties see the significance and benefit of the thing, that is brought through customer value or customer insight.' (A6)

One of the informants (A4) emphasised the professionalism and quality of work.

> *'Evidence has a crucial role in evaluating the solution's functionality and impact, as well as accomplishing positive change. At best, evidence creates positive attitude towards service design, and serves the professional appreciation of service designers as well, especially in public sector service design projects.'* (A4)

Often in service design projects, success can only be measured after the project is closed. One of the informants (A3) explained that this requires a motivated and committed owner within the organisation to achieve the agreed goals and the largest positive impact.

> *'A successful project always demand ownership. That there is someone who owns the project, pushes it through, and is strongly committed in it. If you don't have that, often developmental projects don't lead to anywhere after we detach from them. It can be a sort of momentary momentum, where, of course, new thinking can arise and maybe some momentary motivation. But it quickly returns back to the same rut.'* (A3)

> *'From the service designer's point of view, the greatest success is, when the client, i.e. the public sector organisation has been successfully convinced so, that they understand to change their operations, and that they have to be serious in executing the plan, even though the service design agency is no longer with them.'* (A3)

4.5.2 Evidence as a Support for Decision Making

In the following, the informants tell their experiences about how evidence is used to support decision-making throughout the public service design process.

Most of the informants explained that the decisions should be based on a solid evaluation grounded in the best evidence available. '*Best evidence*' means evidence of which is information from relevant, valid research that has been conducted using appropriate methods – not necessarily the highest or strongest evidence. In '*best evidence*', credibility and believability counts, and one of the informants (A1) described how it can be used to convince top management by creating a sense of urgency.

> *'The leaders have to be excited first, and give permission to do development work. They don't necessarily have to be terribly involved themselves, but they have to understand and give freedom to the others to work on it.'* (A1)

Leading change through complexity requires leadership capabilities at many levels. One of the informants (A6) explained, that for any change to be successful, it requires that all leaders are on board and that they are all in on driving the change from the top, through the rest of the organisation.

> *'Implementing change requires a leadership, that's very pro-development. One criteria for success is, if not the most important one, if the top management is on board or not.'* (A6)

A service project is executed typically to develop a service, which lasts after the project is finished. Regardless of the project size, number of phases, during this time, the project is initiated, planned, prepared, carried out, and closed. One informants (A2) explained that after the project is closed, the organisation

should understand the customer insight work, in order to manage and lead the change.

> *'In a successful project, the organisation starts to take control themselves and starts to move things into the right direction in their own way. Like the organisation is taking the customer insight work seriously and starts procedures. If the result of the service design project is, that the current service isn't working, the organisation understands, they need to fix it to make it better.'* (A2)

One of the informants explained (A5) explained, that in public sector projects, you can often encounter a systemic level.

> *'In many service design projects, especially in the social and health sector, you move between individual and population levels. Sometimes the client organisation can be worried of the thought that we've only talked with a few people - and they feel they can't make a decision on the whole population based on qualitative information. Yes, you can! Sometimes during the process there might have to be decisions made on services regarding the whole population - but in those cases, there can also be used quantitative data, like demographics or other statistical data.'* (A5)

> *'If we think about the hospital as an example. The experiences of the guests, patients and staff reflect the customer viewpoint and lived experience. When it's about things concerning the whole nation, we have to take every experience that we've gathered and then understand the problem from the whole population's viewpoint. If only four people feel the problem is real, then from the whole population's point of view, it probably doesn't concern the whole population, which can be a big and expensive risk in the healthcare environment. A small*

amount does not mean, that the problem can't be fixed because of high costs, but when the issue is examined by combining quantitative and qualitative data, the impact of the issue can be understood on a level concerning the whole population. This, in turn, relaxes the decision makers, because they can make more justified decisions when deciding on the budget' (A5)

5
CONCLUSIONS

The goal of this study has been to find answers to the following research question:

- **What are the evidence-based practices in public service design?**

The following sub-questions elaborate on it:

- **What constitutes evidence in public service design?**
- **How is evidence synthesised and used to inform decision-making?**
- **What is the role of evidence in the different phases of the design process?**

In order to fully answer the main research question and the sub-questions, the following question was a starting point for my empirical work:

- **What is the definition of '*evidence*'?**

In order to answer the main research question, semi-structured interviews with practicing service designers were conducted. The definition of evidence and evidence-based practices was formed with the help of a theoretical framework, that was deepened on the basis of the material collected within the empirical part of the study. From this, several conclusions were derived from the study, concerning the role and the utilisation of evidence in public service design.

In the following, answers to the research questions are presented and summarised. At the end there are also some thoughts on topics where further research is needed.

What constitutes evidence in public service design?

This study has attempted to understand and identify the use of evidence to support service design projects in the public sector, by interviewing practicing service designers. The study confirmed the pre-supposition, that the public sector is a very complex ecosystem. For this reason, also the service design projects differ from each other within the public sector, in terms of scope, goals, complexity and context. Also evidence can be defined in the context of public sector service design, in many different ways, because it can be used in different situations case by case.

The research showed, that evidence can mean in the context of public sector service design many different things, and it can be utilised within the process. The interviewed service designers explained, how they experiment, measure, evaluate and justify their choices within the service design process, as well as how they utilise evidence to e.g. encourage the organisation's strategy work. The service designers also described their understandings of the role and definition of evidence, the characteristics that can be attributed to evidence, the difficulties in its collection and generation of evidence, and suggestions of how to utilise evidence better in design practice.

The following is a summary of three (3) different definitions of evidence, that were derived from the interviews and themes discussed in the findings: 1) *Argumentative Evidence*; 2) *Experience-based Evidence* (or *Qualitative Evidence*); and 3) *Promising Evidence* (or *Potential Evidence*). These forms of evidence can help service designers to add further rigour to the design process, understand how they can better deliver value to the customers through qualitative and experimental methods, as well as stay away from assumptions and pitfalls.

1) Argumentative Evidence

In the context of this study, **ARGUMENTATIVE EVIDENCE** is the most basic form of evidence that allows service designer to show rigour in regard to design practice – and at its simplest form demonstrates the level of rigour in the service design process. Argumentative evidence shows the rigour behind design practice that demonstrates how the conclusions and decisions were made to justify solutions and actions. Therefore, argumentative evidence closely relate to the client perception of the design process as well as the expert status of the service designer. Argumentative evidence can be any convincing evidence that can be used by service designer to support a claim, theory or conclusion – building confidence and showing that the design process was rigorous, thorough and reliable.

The study showed that service designers are required to demonstrate evidence in one form or another, and that is particularly evident in the public service. And particularly so when using qualitative data. The buying organisations are looking for transparency in the decision-making process, as they are required to be transparent and accountable. Therefore, service designers are expected to demonstrate thoroughly e.g. how the data was collected and synthesised, how the decisions were made, what are the arguments and what is the rationale. In this way the buying organisation can be confident about pushing the project through and spending public money on it. Having evidence in design practice and demonstrating the rigorous practice is simply good practice.

2) Experience-based Evidence (or Qualitative Evidence)

In the beginning of the service design process, evidence can be described as **EXPERIENCE-BASED EVIDENCE**, or **QUALITATIVE EVIDENCE**. Unlike anecdotal evidence – that is based on single person's experience – experience-based evidence is drawn from the lived experience of multiple people. Therefore, obtaining a sufficient sample size provides rigour and validity to be reasonably sure of detecting and addressing the real problem. What is considered a sufficient sample size can be entirely subjective, as it relates to the effectiveness of the sample and varies on a case-to-case basis. In service design process, rigour comes from repeated interaction: as the process goes forward and constantly engages with people, it is also building a strong qualitative evidence base.

Since service design is a practice-based approach based on qualitative reasoning, experience-based evidence is typically derived from the understanding and interpreting of the best available research data concerning the lived experience. This data is primarily qualitative in nature and variable quality, that can be collected by means of interviews, reflections and observations, e.g. from citizens, users, customers, as well as frontline staff and service providers that are experiencing or are concerned with the particular problem. Experience-based evidence can also be collected and generated through co-design activities such as participatory design, design games, and artifacts such as cultural probes – that provide a way of collecting data on a specific demographic groups.

Experience-based evidence is relevant when defining a problem or developing an approach to a problem. Experience-based evidence can be used to support the existence or the importance of the problem, the difficulty of the problem or the specific cause of that particular problem. Thus, information becomes evidence when it is used by a service designer to make a synthesis or conclusion regarding the particular matter to which evidence is relevant. The challenging part here is to evaluate the relevant pieces of evidence and distinguish them between facts and opinions.

As service designers collect and analyse user data, they eventually become experts of the lived experience, and are in a position to translate the rich experiences of people in an understandable manner for the benefit of buying organisation, helping them to gain an empathic understanding of people's needs, values and motivations, and also the urgency of the matter. Experience-based evidence can be used to convince the decision-makers in order to get the necessary procedures done to make the implementation of the solution possible, or realise challenging changes. In order to convince the buying organisation and develop a holistic understanding of the problem, service designer can provide different types of evidence e.g. facts, quotations, paraphrases, photographs and videos serve as an evidence to represent people with lived experience, and describe the lived experience of the problem.

3) Promising Evidence (or Potential Evidence)

When making decisions regarding potential solutions to address the particular problem or opportunity, evidence can be described as **PROMISING EVIDENCE**, or **POTENTIAL EVIDENCE**. Promising evidence can be used support the potential solution to a particular problem, and also demonstrate the effectiveness, feasibility or appropriateness of the potential solution. Promising evidence can be used to inform decision-making to ensure that that decisions are well-informed by a good understanding of the likely positive impact of different solutions. It must be noted that promising evidence is not the same thing as empirical evidence – that validates the truth of the solution per se – but it is rather used to justifying the relevance of the potential solution to the context.

Promising evidence can be generated and derived from a variety of experimental studies, such as field experiments, interventions, proof-of-concept prototypes (POC), minimal viable products (MVP) and minimum valuable services (MVS) that can all be tested with real users in real situations. The study showed that while experimentation methods in service design are widely known and used, the application of these methods in design practice is less well documented, especially when it comes to validation of potential solutions, design of interventions, and the rigour of their evaluation methodology. Promising evidence can validate, falsify or invalidate the assumptions underlying the potential solution, and point the right direction.

Although the fidelity of these experiments can vary from low to high depending on the design phase – experiments and interventions designed to test hypotheses should always have measurable indicators to measure impact when possible. Evaluating hypotheses is an integral component of experimental studies, and also an integral part of the learning process. Potential evidence can also validate hypotheses.

How is evidence collected and used to inform decision-making?

The findings from the study have led to the conclusion that service designers utilise a wide variety of different qualitative methods in public service design projects to collect, generate and evaluate evidence. This evidence can be defined as experience-based evidence or qualitative evidence. Because service design is about working with people and designing the way people interact with services, therefore the use of qualitative methods is justified. I would argue that qualitative methods are the only way to gather customer insights and collect information about the specific groups of people that are concerned by the specific problem. There are also many things that can't be achieved through other means – like information on the scope of the challenges or the effectiveness of the potential solutions. The study also showed, that the mere gathering of qualitative data can play a significant part in decision-making process and implementing change, identifying problems and solutions, validating dubious assumptions and bringing other important issues into light.

Promising evidence can be collected and generated through experimentation and experimental designs such as experiments, interventions, prototypes and mini-pilots that can be conducted on real users in a real environment. Experimentation relates to testing the validity of assumptions and hypotheses, as well as evaluating ideas and different potential solutions. Every experiment should be carefully planned, defined and confined, considering the implications both in short and long term. When the right kind of indicators and metrics are attached to the experiments, it is easy to determine if the solution works. And if it does not work, it's known, where it went wrong and what should be developed. Thus, promising evidence can be used to support the design process and decision-making process.

However, in order to support decision-making process, the study showed that planning and evaluation of the experiments need to be taken into consideration. Further actions should not be taken before the thoughtful evaluation of promising evidence has been completed. Indicators and metrics need to be relevant to the context need to be set in advance and aligned to measure the results. They have to also be practical as well as suitable for measuring the experience, i.e. they have to be able to measure what is wanted to measure. The improvement of customer experience can be compared through different kinds of qualitative metrics, that measure e.g. customer satisfaction by comparing the end result to the initial situation. When measuring service experience, the core factor in value creation can for example be customer experience in specific service situations. The research showed that the need for more substantial evidence comes into question, when e.g. local solutions are scaled up to the national or regional level.

The further along the process is going, the more significant the measuring of impact, with quantitative and financial metrics, becomes. Using quantitative metrics can be challenging, because reliable metrics don't always exist. In the context of public sector, quantitative metrics are those based on quantity and endurance, like saving time, speeding up processes, more effective use of space or saving costs. When the change takes place swiftly also its impact can be proven easily, but in complex problems it is harder to see and or takes longer time. Especially when it comes to wicked problems, it can be difficult to pinpoint the causal relationships of the solutions, based on the concrete observations and metrics. In those cases, change and development might be needed to investigate further for a longer time period, even several years through longitudinal studies. With these problems, the statistical differences can be challenging to point out, in which cases selecting the right metrics from the start is crucial in being able to prove the impact.

What is the role of evidence in the different phases of the design process?

The study showed, that service designers had considerably similar views among themselves on the role of evidence and its significance in different stages of the process. The informants often talked about the same things in different terms, depending on their point of view or context. Considerable differences in views were largely explained to the profiles and occupational backgrounds of the informants.

The study showed, that evidence is used through the distinctive stages of the entire service design process; in the beginning qualitative evidence is collected, that characterises the requirement for change, the scope of the problem and the solutions. In the implementation phase, the actual execution phase of the service and at the end of the project, it can be about measuring usability, scaling or measuring impact.

By consistently collecting evidence, the quality of both the solutions and the service being built, can be assured during the process. The study showed, that qualitative methods and development through experimentation are crucial to gather high-quality evidence, to validate solutions and support decision-making. The strength of the evidence needed can depend on the scope and context of the problem. The study showed, that collecting the evidence is critical also to government organisation strategy. Evidence can help define those guidelines on how an organisation can survive its basic task - producing public services.

6
RECOMMENDATIONS

At the end of the research, there are recommendations presented, that can help the process of the service designers working in the public sector service design and their work in practice. The recommendations are about ways to develop by trying out things, as well as utilising quantitative methods in a reasonable way in conjunction with the qualitative methods.

Transparency of the Public Service Design Process

The way in which service design is done in the public sector at the moment, is mainly through public competitive tendering. The actual design process and work is mainly done behind closed doors, which is not useful, when considering the social impact and the transference of best practices. By multiplying best practices and ways of working, larger implications and more desired results would be gained.

Also the client organisations in the public sector could be more transparent about the evidence collected within the different projects, field test results, and communicate about the process and results gained, clearly and openly. This could help strengthen the trust towards service design in the public sector, and create the justification for utilising best practices in possibly other projects and public sector organisations. Especially, if a project is foreseen to have a positive social impact, it would be useful, if these observations could be utilised also in other contexts.

When the process is planned, executed and measured efficiently, it can be suggested, that the service being built, will probably have a positive effect.

- *If there's been useful or promising evidence collected with qualitative methods, the results derived from the experiences of the employees and the customers can be applied to another context. Who else could use these observations and how could they utilise these observations?*

- *How could service design consultancies make sure that they could tell about their process openly, so that best practices would spread also outside the closed doors?*

- *Who should do the impact assessment?*

A typical decision making process, done behind closed doors, doesn't necessarily provide understanding about what the actual impact of the project would have been, if the process would've been done as a transparent process, which has a clear connection to a larger system, or other public sector organisations, that could utilise directly or indirectly the observations or evidence gotten from the project. This kind of co-operation that crosses the governmental branches and organisations could be useful to the whole public sector, leading to better services and results.

Impact Assessment of Public Service Design

It would be recommended to do a research of its own regarding the impact of service design in the public sector, for example by analysing completed projects using a comparative setting, longitudinal study, or before-and-after study. Knowledge of a relevant case studies is possible to collect, but it can be very laborious. Also, impact is not always straightforward, and collecting the information and digesting it can be challenging.

Collecting Best Practices in the Public Sector

In addition to the impact assessment, there should also be a collection of best practices in public service design. At the moment, there's no way to collect, analyse and produce best practice cases. Service design consultancies could better document and follow up on their own work and actions within the public sector, as well as strive for transparency so that also others could utilise the vast knowledge that's accumulated during the project, and thus help with their own actions to bring more professionalism into service design and further the use of evidence-based practices.

The Importance of Indicators and Metrics

The study showed, that different kinds of indicators and metrics are important tools for collecting the evidence, and that the quality and strength of the evidence are highly dependant on the nature, scope and goals of the service design project. However, sometimes there are situations when the desired impact can be difficult to measure, or can be too complicated to be transformed into something quantifiable. It must be also noted that when potential solutions are identified, tested and developed, there are also situations where more clear and convincing evidence is needed. For example, when it's about interventions or service procedures concerning people's health and wellbeing, it has to be made sure that the solution does not risk anyone's health or safety. In these cases, qualitative research data can be backed up by quantitative indicators e.g. using demographic information or statistical data that's relevant to the project.

LIST OF TABLES

LIST OF FIGURES

REFERENCES

Achinstein, P. (1983). *The Concept of Evidence*. Oxford University Press.

Achinstein, P. (2001). *The Book of Evidence*, p.5. Oxford Studies in Philosophy of Science. New York: Oxford University Press.

Andriopoulos, C. and Lowe, A. (2000). *Enhancing Organisational Creativity: The Process of Perpetual Challenging*. Management Decision 38.10, pp. 734–742.

Argote, L. and Ingram, P. (2000). *Knowledge Transfer: A Basis for Competitive Advantage in Firms*. Organizational Behavior and Human Decision Processes 82.1, pp. 150–169.

Arlander, N. (2014). *Procurement System HILMA as a Boundary Object*. Master's Thesis. Aalto University School of Science. Retrieved 20 Feb. 2017 from <https://aaltodoc.aalto.fi/bitstream/handle/123456789/ 14103/master_Arlander_Nora_2014.pdf?sequence=2&isAllowed=y>

Ball, L. J. and Christensen, B. T. (2009). *Analogical Reasoning and Mental Simulation in Design: Two Strategies Linked to Uncertainty Resolution*. Design Studies, 30(2), 169-186.

Bason, C (2010a). *Co-creation is Key to Innovation in Government*, p. 14-17. Ipsos MORI Understanding Society. Winter 2010. Retrieved 20 Feb. 2017 from <http://mind-lab.dk/wp-content/uploads/ 2014/09/Understanding_Society_Winter_2010_CBason.pdf>

Bason, C. (2010b). *Leading Public Sector Innovation*. Co-Creating for a Better Society. Bristol, United Kingdom: The Policy Press.

Bason C. and Schneider, A. (2015). *Public Design in Global Perspective: Empirical Trends*, p. 23-34. In: C. Bason (ed.), Design For Policy. Design for Social Responsibility Series. Farnham: Gower.

Bazeley, P. and Jackson, K. (2013). *Qualitative Data Analysis with NVivo*, p. 33. Sage Publishing.

Bediko, S. (2015). *Project Oracle – Children & Youth Evidence Hub*. Presentation by Steven Bediko, Chief Executive of Social Innovation Partnership. Evidence Exchange Webinar. 1 Sep. 2015.

Behavioural Insights Team (2012). *Applying Behavioural Insights to Reduce Fraud, Error and Debt*, p. 34. Cabinet Office. Retrieved 20 Feb. 2017 from <https://www.gov.uk/government/publications/fraud-error- and-debt-behavioural-insights-team-paper>

Blomkvist, J. (2011). *Conceptualising Prototypes in Service Design*, p. 30. Thesis No. 101. University of Linköping, Sweden.

Blomkvist, J. and Holmlid, S. (2012). *Service Prototyping According to Service Design Practitioners*. IDA. University of Linköping, Sweden.

Blume, T. (2016). *When Did It Become Trendy to Start Slating Randomised Controlled Trials?* Alliance 4 Useful Evidence. Nesta Retrieved 20 Feb. 2017 from <http://www.alliance4usefulevidence.org/when- did-it-become-trendy-to-start-slating-randomised-controlled-trials/>

Borins, S. (2001). *Encouraging Innovation in the Public Sector*, p. 310-319. Journal of Intellectual Capital, Vol. 2, No. 3.

Bowling A. (2002). *Research Methods in Health: Investigating Health and Health Services*, p. 118. 2nd ed. Open University Press. New York, NY.

Brandt, E. (2006). *Designing Exploratory Design Games: A Framework for Participation in Participatory Design?*, p. 57-66. Proceedings of the Ninth Conference on Participatory Design (PDC 2006). New York: ACM Press.

Brechin, A. and Sidell, M. (2000). *Ways of Knowing*. In: R. Gomm and C. Davies (eds.) Using Evidence in Health Care. Buckingham: Open University Press.

Breckon, J. (2015a). *Using Research Evidence - A Practice Guide*. Nesta. London: Alliance for Useful Evidence.

Breckon, J. (2015b). *What is Good Evidence?* Intro by Jonathan Breckon, Head of Alliance for Useful Evidence. Social Policy Evidence. Evidence Exchange Webinar. 1 Sep. 2015.

Breckon, J. and Dodson, J. (2016). *Using Evidence – What works?*, p. 28. A Discussion Paper. London: Alliance for Useful Evidence. Retrieved 20 Feb. 2017 from <http://www.nesta.org.uk/sites/ default/files/using_evidence_what_works.pdf>

Briggs, M. and Flemming, K. (2007). *Living with leg ulceration: a synthesis of qualitative research.* J Adv Nurs. 2007 Aug; 59(4):319-28. Epub 2007 Jul 2.

Brooks, J. D. (2015). *On Training in Language Documentation and Capacity Building in Papua New Guinea: A Response to Bird et al*. University of California at Santa Barbara. Retrieved 20. Feb 2017 from: <https://scholarspace.manoa.hawaii.edu/bitstream/handle/10125/24629/brooks.pdf>

Bryman, A. (1984). *The Debate About Quantitative and Qualitative Research: A Question of Method or Epistemology?*, p. 75-92. The British Journal of Sociology. Vol. 35. No. 1. The London School of Economics and Political Science.

Bryman, A. (2008). *Social Research Methods*, p. 393. Oxford: Oxford University Press.

Buchanan, R. (2001). *Design and the New Rhetoric: Productive Arts in the Philosophy of Culture*. Philosophy and Rhetoric, 34(3), 183-206.

Buijs, J. (2007). *Innovation Leaders Should Be Controlled Schizophrenics*. Creativity and Innovation Management 16.2, pp. 203–210.

Burns, P. B., Rohrich, R. J., & Chung, K. C. (2011). *The Levels of Evidence and Their Role in Evidence-Based Medicine*. Plastic and Reconstructive Surgery, 128(1), 305-310. Retrieved 20 Feb. 2017 from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3124652/>

Cabinet Office (1999a). *The Modernising Government*. White Paper. Presented to Parliament by the Prime Minister and the Minister for the Cabinet Office by Command of Her Majesty. Retrieved 20 Feb. 2017 from <http://www.archive.official-documents.co.uk/document/cm43/4310/4310-02.htm>

Cabinet Office (1999b). *Professional Policy-Making for the Twenty-first Century*, p. 31. Report by Strategic Policy Making Team. Cabinet Office.

Canadian Task Force on the Periodic Health Examination (1979). *The Periodic Health Examination*, p. 121: 1193–1254. CMAJ.

Carmichael, J. and Kaufmann, D. (2001). *Public Sector Governance and the Finance Sector*, p. 11. Paper prepared for the conference on: Policy Challenges for the Financial Sector in the Context of Globalization, Washington DC, June 2001.

Carr, V. L., Sangriorgi D., Büscher, M. Junginger, S. & Cooper. R. (2011). *Integrating Evidence-based Design and Experience-based Approaches in Healthcare Service Design*, p. 12-14. HERD: Health Environments Research & Design Journal.

Castillo, N. M. & Wagner, D. A. (2013). *Gold Standard? – The Use of Randomized Controlled Trials for International Educational Policy*, p. 166-173. Comparative Education Review. 58(1).

Christiansen, J. (2016). *Embedding Design: Towards Cultural Change in Government*, p. 48-58. In: B. Mager (ed.). Service Design Impact Report: Public Sector. Köln: Service Design Network. Retrieved 20 Feb. 2017 from <https://www.service-design-network.org/books- and-reports/impact-report-public-sector>

Christiansen, J. and Schulman, S. (2015). *Inside or Out? IWD vs. MindLab*. InWithForward. Retrieved 20 Feb. 2017 from <http://inwithforward.com/resources/inside-or-out-iwf-vs-mindlab>

Clinks (2017). *Standards of Evidence Guide*. Retrieved 20 Feb. 2017 from <http://www.clinks.org/sites/ default/files/StandardsofEvidenceGuide.pdf>

Codinhoto, R. (2013). *Evidence and Design: An Investigation of the Use of Evidence in the Design of Healthcare Environments*. School of the Built Environment. College of Science and Technology. The University of Salford, Salford, Uk. Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor of Philosophy, February 2013.

Collins, K. & Cook M–R. (2014). *Ethics of Participation Within Service Design*, p. 38-39. In: D. Sangriorgi et al. Mapping and Developing Service Design Research in the UK, Chapter 4 - Emerging Research Themes. Project Report. Service Design Research UK Network. Lancaster, UK.

Commission of European Communities (2009). *Commission Staff Working Document: Design as a Driver of User-Centred Innovation*. (501). Brussels, Belgium: The Commission of European Communities. Retrieved 20 Feb. 2017 from <http://ec.europa.eu>

Commission of European Communities (2013). *Commission Staff Working Document: Implementing An Action Plan for Design-Driven Innovation*. (380). Brussels, Belgium: The Commission of European Communities. (PDF). Retrieved 20 Feb. 2017 from <http://ec.europa.eu>

Cook, T. D. and Gorard, S. (2007). *What counts and should count as evidence?* In: Evidence in education: Linking research and policy. Paris.

Daly, J., Willis, K., Small R., Green, J., Welch, N., Kealy, M. and Hughes, E. (2007). *A Hierarchy of Evidence for Assessing Qualitative Health Research*, p. 45-46. Journal of Clinical Epidemiology 60. Elsevier.

Davies, H. T. O., & Nutley, S. M. (2008). *Learning More About How Research-based Knowledge Gets Used: Guidance in the Development of New Empirical Research*, p. 2. William T. Grant Foundation, New York, NY.

Demos Helsinki, Avanto Helsinki and Aalto University (2015). *Design for Government: Human-centric Governance Through Experiments*. Prime Minister's Office, Finland. ISBN: 978-952-287-179-4. Retrieved 20 Feb. 2017 from <http://vnk.fi/documents/10616/1456483/DfG-raportti.pdf/ 4072bebd-ff36-4876-89e2-09076d8142df> .

Design Commission (2013). *Restarting Britain 2*, p.31. Retrieved 20 Feb. 2017 from <http://www.policyconnect.org.uk/apdig/research/restarting-britain-2-design-public-services>

Design for Government (2017). Aalto University. Retrieved 20 Feb. 2017 from <http://dfg-course.aalto.fi/2015/about-design-for-government>

Donetto, S., Pierri, P. Tsianakas, V. and Robert, G. (2015). *Experience-based Co-design and Healthcare Improvement: Realizing Participatory Design in the Public Sector*. The Design Journal. An International Journal for All Aspects of Design. Volume 18, 2015. Issue 2. 227-248.

Drew, C. (2016). *Engaging people as Service Users and Citizens*, p. 78-83. In: B. Mager (ed.). Service Design Impact Report: Public Sector. Köln: Service Design Network. Retrieved 20 Feb. 2017 from <https://www.service-design-network.org/books- and-reports/impact-report-public-sector>

Edovald, T. and Firpo, T. (2016). *Running Randomised Controlled Trials In Innovation, Entrepreneurship and Growth: An Introductory Guide*. Innovation Growth Lab and Nesta. Retrieved 20 Feb. 2017 from <http://www.nesta.org.uk/sites/default/files/a_guide_to_rcts_-_igl.pdf>

ERCAS (2015). *Public Integrity and Trust in Europe. European Research Centre for Anticorruption and State-Building* (ERCAS), Hertie School of Governance. Berlin. Retrieved 20 Feb. 2017 from <https://www.government.nl/binaries/government/documents/reports/2016/01/18/public- integrityand-trust-in-europe/hertie-2015-public-integrityand-trust-in-europe-final.pdf>

Eskelinen, S. (2002). *Dialogilla kohti positiivista mielikuvaa. Kuluttajaviraston viestintätutkimuksen loppuraportti.* Kauppa- ja teollisuusministeriön tutkimuksia ja raportteja. Helsinki: Edita.

European Commission (2005). *Explaratory Note – Framework Agreements – Classic Directive*. Public Procurement Policy. Retrieved 20 Feb. 2017 from <http://ec.europa.eu/internal_market/ publicprocurement/docs/explan-notes/classic-dir-framework_en.pdf>

Feast, L. (2012). *Professional Perspectives on Collaborative Design Work*. CoDesign, 8(4), 215-230. Feast, L. (2015). Investigating Collaboration in Interdisciplinary Design Teams. Swinburne University of Technology, Australia.

Fuda, P. (2009). *Why Change Efforts Fail*. White Paper. The Alignment Partnership (TAP). Retrieved 20 Feb. 2017 from <http://www.tap.net.au/resources/TAP_ WhitePaper_ WhyChangeEffortsFail.pdf

Gage, M. (2002). *Making Emotional Connections Through Participatory Design.* Boxes and Arrows. Retrieved 20 Feb. 2017 from <http://boxesandarrows.com/making-emotional-connections-through- participatory-design/>

Glasby, J., Walshe, K., and Harvey, G. (2007). *Making Evidence Fit for purpose in Decision Making: A Case Study of the Hospital Discharge of Older People*. Evidence and Policy, 3(3), 425–437.

Goodman, S. N. and Royall, R. (1988). *Evidence and Scientific Research*. American Journal of Public Health 78, 1568-1574.

Gross, L. and Smith, K.L. (2015). *Applying Design Thinking to Citizen Service*. Retrieved 20 Feb. 2017 from <https://www.td.org/Publications/Magazines/The-Public-Manager/Archives/2015/Summer/ Applying-Design-Thinking-to-Citizen-Service>

Hagen, P., Nicholas, M. & Rahilly K. (2014). *Integrating UX and Evidence-based Approaches to Design Effective Youth Mental Health Services*, p. 22-25. UX New Zealand. Retrieved 20 Feb. 2017 from <http://www.smallfire.co.nz/2014/01/25/integrating-user-experience-and-evidence-based- approaches-to-design>

Hartley, J. (2005). *Innovation in Governance and Public Services: Past and Present*, p. 27-34. Public Money And Management. Vol. 25, No. 1.

Helsinki Design Weekly (2016). *The First Chief Design Officer of Helsinki is Anne Stenros*. Helsinki Design Week. 12. July 2016. Retrieved 20 Feb. 2017 from: <http://www.helsinkidesignweek.com/weekly/ the-first-chief-design-officer-of-helsinki- is-anne-stenros/?lang=en>

HM Treasury (2011). *The Magenta Book: Guidance for Evaluation*. London: HM Treasury.

Hodge, G. A. and Coghill, K. (2004). *Governing the Privatised State: The Accountability Challenge*, p. 3. Working Paper 57/04. October 2004. Monash University. This paper was presented at the International Research Symposium on Public Management (IRSPM) VII, Hong Kong, April, 2003

Humphreys, P. C. (1998). *Improving Public Service Delivery*. Committee for Public Management Research. Dublin: Institute of Public Administration.

IDEO, Design for Europe and Nesta (2016). *Designing for Public Services: A Practical Guide*. Retrieved 20 Feb. 2017 from <http://designforeurope.eu/sites/default/files/asset/document/ Nesta_Ideo_Guide_021216_0.pdf>

Ilic, M. and Puttick, R. (2012). *The Development of Project Oracle – Generating and Using Evidence In The Real World*. London: Mayor of London and Nesta. Retrieved 20 Feb. 2017 from <https://www.nesta.org.uk/sites/default/files/the_development_of_project_oracle.pdf>

The Interaction Design Foundation (2017). *The Principles of Service Design Thinking: Building Better Services*. Retrieved 20 Feb. 2017 from <https://www.interaction-design.org/literature/article/ the-principles-of-service-design-thinking-building-better-services>

International Telecommunications Union (2016). *ICT Facts and Figures 2016*. Retrieved 20 Feb. 2017 from <https://www.itu.int>

Johnson, R. N. and Libecap G. D. (1994. *The Implications of a Protected Bureaucracy*, p. 154-176. In. R. N. Johnson and G. D. Libecap (eds.). The Federal Civil Service System and The Problem of Bureaucracy. University of Chicago Press.

Kaiser, K. (2009). *Protecting Respondent Confidentiality in Qualitative Research*. Qualitative Health Research, 19(11). 1632-1641.

Kirchberger, E–M. & Tether, B. (2014). *Should Service Design Professionalize, and How?*, p. 40-41. In: D. Sangriorgi et al. Mapping and Developing Service Design Research in the UK, Chapter 4 - Emerging Research Themes. Project Report. Service Design Research UK Network. Lancaster, UK.

Kirk, J., and Miller, M. L. (1986). *Reliability and Validity in Qualitative Research*. Thousand Oaks, CA: Sage Publications.

Kvernbekk, T. (2008). *Evidence-based Practice in Education: Functions of Evidence and Causal Presuppositions.* Volume 147 of Routledge Research in Education. Routledge.

Laboratorio de Gobierno (2017). Website. Retrieved 20. Feb 2017 from <https://lab.gob.cl>

Lee, A. S. and Hubona, G. S. (2009). *A Scientific Basis for Rigor in Information Systems Research*, p. 237-262. MIS Quarterly (33:2).

Leinonkoski, H. (2013). *Service Design - An Approach to Better Public Services? A Civil Servant's View*, p. 14-15. London: The Finnish Institute In London. Retrieved 20 Feb. 2017 from <http://www.fininst.uk/en/articles/540-new-working-paper-published-on-service-design>

Lincoln, Y. and Guba, E. (1985). *Naturalistic Inquiry.* Vol. 75. Beverly Hills: Sage Publications.

Lincoln, Y. and Guba, E. (2005). *Paradigmatic Controversies, Contradictions, and Emerging Confluences.* In: N. Denzin and Y. Lincoln (eds.). The Sage Handbook of Qualitative Research. (3rd ed., pp. 191-216). Thousand Oaks, CA: Sage Publication.

Loosemore, T. (2015). *Digital Government: Not Complicated*. Just Hard. Presentation by Tom Loosemore, former director of UK Digital Services. Helsinki Art Museum (HAM). 11 Dec. 2015.

Mager, B. (2016a). *Innovating Public Services*. In: B. Mager (ed.). Service Design Impact Report: Public Sector, p. 8, 12-13. Köln: Service Design Network. Retrieved 20 Feb. 2017 from <https://www.service-design-network.org/books-and-reports/impact-report-public-sector>

Mager, B. (ed.) (2016b). *Service Design Impact Report: Public Sector*. Köln: Service Design Network. Retrieved 20 Feb. 2017 from <https://www.service-design-network.org/books- and-reports/impact-report-public-sector>

Marston, C., and King, E. (2006). *Factors that shape young people's sexual behaviour: a systematic review*. Lancet. 2006 Nov 4;368(9547):1581-6.

Maxwell, J. A. (1992). *Understanding and Validity in Qualitative Research*, p. 279-300. Harvard Educational Review (62:3).

Mazé, R. (2016). *Design for Government (DfG) 2016 Final Show Opening. Introductory remarks by Professor Ramia Mazé*. DfG is advanced studio course in Aalto University's Creative Sustainability Master's Programme. Retrieved 20 Feb. 2017 from <https://vimeo.com/174681824>

McLoughlin, C. (2015). *When Does Service Delivery Improve the Legitimacy of a Fragile or Conflict-affected State?*, p. 343. University of Birmingham. Governance: An International Journal of Policy, Administration and Institutions, Vol. 28, N. 3, July 2015 (pp. 341-356).

Meidutē, I. & Paliulis, N. (2011). *Feasibility Study of Public-private Partnership*, p. 257-271. International Journal of Strategic Property Management. 15(3). Business Source Complete. Viewed 14 Dec. 2015 at EBSCOhost.

Merriam-Webster (2017). *Evidence*. Merriam-Webster Incorporated. Retrieved 20 Feb. 2017 from <https://www.merriam-webster.com/dictionary/evidence>

Miles, M. and Huberman, M. A. (1984). *Qualitative Data Analysis*. California: Sage Publishing.

Miller, D. B. and Rudnick, M. (2012). *A Framework Document for Evidence-Based Programme Design on Reintegration*. Produced by the United Nations Institute for Disarmament Research in support of the Inter-Agency Working Group on Disarmament, Demobilisation and Reintegration. UNIDIR. United Nations Institute for Disarmament Research. Geneva, Switzerland.

Miller, S. and Fredericks, M. (2003). *The Nature of 'Evidence' in Qualitative Research Methods*, p. 2. International Journal of Qualitative Methods, 2(1). Article 4. Retrieved 20 Feb. 2017 from <http://www.ualberta.ca/~iiqm/backissues/2_1/html/miller.html>

Miller, S. I. and Safer, L. A. (1993). *Evidence, Ethics and Social Policy Dilemmas*. Education Policy Analysis Archives. 1(9). Retrieved 20 Feb. 2017 from <http://epaa.asu.edu/ojs/article/view/686>

Milliken, J. and Krause K. (2002). *State Failure, State Collapse, and State Reconstruction: Concepts, Lessons and Strategies*, p. 760. Institute of Social Studies. Development and Change 33(5): 753–774. Oxford: Blackwell Publishers.

Mimica, N. (2013). *Applying Behavioural Insights to Policy-Making: Results, Promises and Limitations*. Commissioner for Consumer Policy. 2013. Retrieved 20 Feb. 2017 from <http://europa.eu/rapid/ press-release_SPEECH-13-761_en.htm>

Mohr, L. B. (1969). *Determinants of Innovation in Organizations*, p. 111-126. The American Political Science Review. Vol. 63, No. 1.

Moisander, J. and Valtonen, A. (2006). *Qualitative Marketing Research: A Cultural Approach*, p. 21-28. London: Sage Publishing.

Morgan, G. A., Gliner, J. A. and Harmon, R. J. (2000). *Internal Validity.* Journal of the American Academy of Child and Adolescent Psychiatry, 39, 529-531.

Moskal, B. M., Leydens, J. A. and Pavlich, M. J. (2002). *Validity, Reliability and the Assessment of Engineering Education*, p. 263-365. Journal of Engineering Education. Vol. 91, Issue 3.

Muir, R. and Parker, I. (2014). *Many to Many.* Institute for Public Policy Research. London. Retrieved 20 Feb. 2017 from <http://www.ippr.org/files/images/media/files/publication/2014/02/ Many-to-many_Feb2014_11865.pdf>

Mulgan, G. and Leadbeater, C. (2013). *Systems Innovation Discussion Paper*, p. 7-9, 19. London: Nesta. (PDF). Retrieved 20 Feb. 2017 from <http://www.nesta.org.uk/sites/default/files/ systems_innovation_discussion_paper.pdf>

Mullen, E. J. (2002). *Evidence-based Knowledge: Designs for Enhancing Practitioner Use of Research Findings*. Paper presented at the 4th International Conference on Evaluation for Practice, University of Tampere, Tampere, Finland.

Mullen, E. J. (2004). *Evidence-based Practice in a Social Work Context – the United States Case*, p. 54. National Research and Development Centre for Welfare and Health (STAKES), FinSoc Working Papers 2/2004.

Mumford, M. D., Scott, G. M., Gaddis, B., and Strange, J. M. (2002). *Leading Creative People: Orchestrating Expertise and Relationships*. The Leadership Quarterly 13.6, pp. 705–750.

Munro, S. A., Lewin, S. A., Smith, H. J., Engel, M. E., Fretheim, A. and Volmink, J. (2007). *Patient Adherence to Tuberculosis Treatment: A Systematic Review of Qualitative Research*. PLOS Medicine. 2007 Jul 24; 4(7):e238.

Murray, C.H. (1990). *The Civil Service Observed,* p. 151. Dublin: Institute of Public Administration.

Murray, R., Caulier-Grice, J. and Mulgan, G. (2010). *The Open Book of Social Innovation, p. 91. Social Innovator Series: Ways to Design, Develop and Grow Social Innovation*. The Young Foundation and Nesta. Retrieved 20 Feb. 2017 from <http://youngfoundation.org/wp-content/uploads/2012/10/ The-Open-Book-of-Social-Innovationg.pdf>

Myers, M. D. (2013). *Qualitative Research in Business and Management*, p. 123. Sage Publishing.

Myers, M. D. and Newman, M. (2007). *The Qualitative Interview in IS research: Examining The Craft*, p. 2-26. Information and Organization, Vol. 17, No. 1. January 2007.

Ní Ógáin, E. (2015). *Impact Measurement in Impact Investing – Learning From Practice*. London: Nesta Impact Investments. Retrieved 20 Feb. 2017 from <http://www.nesta.org.uk/sites/default/files/ impact_measurement_in_impact_investing.pdf>

Noyes, J. and Popay, J. (2007). *Directly observed therapy and tuberculosis: how can a systematic review of qualitative research contribute to improving services? A qualitative meta-synthesis*. J Adv Nurs. 2007 Feb; 57(3). 227-43.

Nutley, S., Powell, A. and Davies, H. (2013). *What Counts As Good Evidence?* Provocation Paper for the Alliance for Useful Evidence. February 2013. London: Alliance for Useful Evidence.

O'Reilly, B. (2016). *How to Implement Hypothesis-Driven Development*. ThoughtWorks Blog. Retrieved 20 Feb. 2017 from <https://www.thoughtworks.com/insights/blog/how-implement-hypothesis- driven-development>

Oancea, A. and Pring, R. (2009). *The Importance of Being Thorough: On Systematic Accumulations of 'What Works' in Education Research*, p. 11-35. In: D. Bridges et al. (eds.). Evidence-Based Educational Policy: What Evidence? What Basis? Whose Policy. New Jersey: John Wiley & Sons.

Ohala, J. I. (1986). *Consumer's Guide to Evidence in Phonology*, p. 3-26. Phonology Yearbook 3. University of California. Berkeley.

Oxford Compact English Dictionary (1971) 10th ed. Oxford University Press: Oxford.

Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods*, p. 227; 230; 341-343. Los Angeles, CA: Sage.

Pawson, R. (2012). *Evidence-based Policy. A Realist Perspective*. Los Angeles: Sage Publishing.

Pearce, S. (2000). *Differences Between Private and Public Sector IT*, p. 4. In: Government IT Projects. Report 200, July 2003. London: Parliamentary Office of Science and Technology. Retrieved 20 Feb. 2017 from <http://www.parliament.uk/documents/post/pr200.pdf>

Popper, K. (1959). *The Logic of Scientific Discovery*. Basic Books. New York: NY.

Potter, C. and Brough, R. (2004). *Systemic Capacity Building: A Hierarchy of Needs*. Health Policy Planning. 19(5): 336-345.

Project Oracle (2017). *Standards of Evidence*. Retrieved 20 Feb. 2017 from <http://project-oracle.com/projects/standards-of-evidence/>

Puttick, R. and Ludlow, J. (2012). *Standards of Evidence for Impact Investing*. London: Nesta.

Ramírez, R. and Mannervik, U. (2008). *Designing Value-creating Systems*. In: L. Kimbell and V. P. Seidel (eds.). Designing for Services. Said Business School. Oxford: University of Oxford.

Rebolledo, N. (2016). *The Value of Service Design in Policy Making*, p. 40-46. In: B. Mager (ed.). Service Design Impact Report: Public Sector. Köln: Service Design Network. Retrieved 20 Feb. 2017 from <https://www.service-design-network.org/books- and-reports/impact-report-public-sector>

Reiss, J. (2014). *What's Wrong With Our Theories of Evidence?*. THEORIA: An International Journal for Theory, History and Foundations of Science 29(2): 283-306

Rentoul, J. (2013). *Reinventing Blairism*. Retrieved 14 Jan. 2016 from <http://blogs.independent.co.uk/ 2013/03/09/reinventing-blairism>

Republic of Estonia (2014). *e-Estonia. Fact Sheet.* August 2014. Republic of Estonia. Ministry of Foreign Affairs. Retrieved 20 Feb. 2017 from: <http://vm.ee/sites/default/files/elfinder/article_files/e-estonia.pdf>

ResearchGate (2017). *Is it possible to set hypothesis for qualitative data? Discussion.* Retrieved 20 Feb. 2017 from <https://www.researchgate.net/post/is_it_possible_to_set_hypothesis_for_qualitative_data>

Ridenour, C. S. and Newman, I. (2008). *Mixed Methods Research: Exploring the Interactive Continuum*. Carbondale, IL: Southern Illinois University Press.

Ries, E. (2011). *The Lean Startup: How Today's Entrepreneurs use Continuous Innovation to Create Radically Successful Businesses*. New York: Crown Business.

Ross, T. (2012). *Chapter 1: Evidence-based Methods*. In: T. Ross (2012). A Survival Guide for Health Research Methods. 1st Edition. Open University Press.

Rotberg, R. I. (2014). *When States Fail: Causes and Consequences*, p. 2-3. Princeton: Princeton University Press.

Rutter, T. (2015). *The Rise of Nudge – The Unit Helping Politicians to Fathom Human Behaviour.* The Guardian. Retrieved 14 Jan. 2016 from <http://www.theguardian.com/public-leaders-network/ 2015/jul/23/rise-nudge-unit-politicians-human-behaviour>

Rusonis S. (2015). *How to Write a Solid A/B Test Hypothesis*. Optimizely. Retrieved 20 Feb. 2017 from <https://blog.optimizely.com/2015/01/29/why-an-experiment-without-a-hypothesis-is-dead-on-arrival/>

Sackett, D. L. (1989). *Rules of Evidence and Clinical Recommendations on the Use of Antithrombotic Agents*, p. 95; 2S–4S. Chest.

Sackett, D. L., Rosenberg, W.M., Gray J.A., Haynes, R.B. and Richardson, W.S. (1996). *Evidence based Medicine: What it is and what it isn't – It's about integrating individual clinical expertise and the best external evidence*, p. 71. BMJ. Volume 312.

Salmelin, A. (2014). *Successful Service Development in the Public Sector: A Case Study of a Service Design Project Carried Out in Short-term Care Services in the Town of Turku*, p. 77. Master's Thesis. Laurea University of Applied Sciences. Retrieved 20 Feb. 2017 from <https://www.theseus.fi/ bitstream/handle/10024/78590/ Annaliisa_Salmelin_Master_Thesis_2014.pdf>

Sangriorgi, D. and Pacenti, E. (2008). *Emerging Practices of Service Design Driven Innovation*, p. 1. SDN Network Amsterdam. November 2008. This text builds on a research proposal 'Service Design and Innovation: The Future of Service Design Research' (Linköping University).

Sangriorgi D., Prendiville A. & Ricketts, A. (2014). *Mapping and Developing Service Design Research in the UK*, p. 3-29. Project Report. Service Design Research UK Network. Lancaster, UK.

Scott, M. (2014). *Estonians Embrace Life In A Digital World*. The New York Times. Retrieved 20 Feb. 2017 from <https://www.nytimes.com/2014/10/09/business/international/ estonians-embrace-life-in-a-digital-world.html>

SEE Platform (2013). *Design for Public Good*, p.8. Design Council, Design Wales, Danish Design Centre and Aalto University. Sharing Experience Europe. Policy Innovation Design. Retrieved 20 Feb. 2017 from <http://www.seeplatform.eu/docs/Design%20For%20Public%20Good%20May%202013.pdf>

Service Design Network (2016). *Online survey conducted by Service Design Network (SDN) from July to August, 2016*, p. 13. In: B. Mager (ed.), Service Design Impact Report: Public Sector. Köln: Service Design Network. (PDF). Retrieved 20 Feb. 2017 from <https://www.service-design-network.org/books- and-reports/impact-report-public-sector>

Service Design Network (2017). *Service Design Network (SDN) Website*. Retrieved 20 Feb. 2017 from <https://www.service-design-network.org/ organisations/service-design-network>

Sevaldson, B. (2011). *GIGA-Mapping: Visualisation for Complexity and Systems Thinking in Design*. Oslo School of Architecture and Design. Ocean Design Research Association. Paper presented at the Nordic Design Research Conference.

Snook and DMA (2014). *Service Design Principles for Working with the Public Sector: A Collaborative Think Piece*, p. 4-12. Design Managers Australia. Retrieved 20 Feb. 2017 from <http://designmanagers.com.au/wp-content/uploads/2014/06/dma_snook_article.pdf>

Skoufias, E. and McClafferty, B. (2001). *Is Progresa Working? Summary of the Results of an Evaluation by IFPRI*. Food Consumption and Nutrition Division. International Food Policy Research Institute. July 2001. Retrieved 20 Feb. 2017 from: <http://ageconsearch.umn.edu/bitstream/16418/1/fc010118.pdf>

Social Innovator (2017). *In-house Innovation Units. Social Exchange*. The Young Foundation and Nesta and Retrieved 20 Feb. 2017 from <http://www.socialinnovator.info/connecting-people-ideas-and- resources/innovation-intermediaries/teams/house-innovation-units>

Staley, K. W. (2004). *The Evidence for the Top Quark: Objectivity and Bias in Collaborative Experimentation*. Cambridge University Press.

Stenbacka, C. (2001). *Qualitative Research Requires Quality Concepts of its Own*, p. 551-555. Management Decision (39:7).

Sustar, H. & Feast, L. (2016). *Problematizing Evidence-Based Design: A Case Study of Designing for Services in the Finnish Government*. Proceedings of DRS 2016, Design Research Society 50th Anniversary Conference. Brighton, UK, 27–30 June 2016.

Taylor, M. C. and Savin-Baden, M. (2001). *Whose 'evidence' are we applying?* British Journal of Occupational Therapy, 64, 213.

Tenders Electronic Daily (2017). *Tenders Electronic Daily (TED) Website*. Supplement to the Official Journal of the EU. Retrieved 20 Feb. 2017 from <http://ted.europa.eu>

Thaler, R. H. and Sunstein, C. R. (2009). *Nudge: Improving Decisions About Health, Wealth and Happiness*. New York: Penguin Books.

Thomke, S. (2003). *R&D Comes to Services*. Bank of America's Pathbreaking Experiments. Harvard Business Review 81.4, pp. 70–9.

Thomke, S., Von Hippel, E. and Franke, R. (1998). *Modes of Experimentation: An Innovation Process and Competitive Variable*. Research Policy 27.3, pp. 315–332.

Thomke, S. and Reinertsen, D. (1998). *Agile Product Development: Managing Development Flexibility in Uncertain Environments*. In: California Management Review 41.1, pp. 8–30.

Torres, T. (2017). *Take Ten: The Ladder of Evidence*. Product talk. Retrieved 8 Mar. 2017 from <https://www.producttalk.org/2017/03/ladder-of-evidence/>

Trevithick, E. (2015). *The Standards of Evidence: Our Next Move*. The Social Innovation Partnership. Retrieved 20 Feb. 2017 from <https://tsip.co.uk/news/2015/the-standards-of-evidence-our-next-move>

Tuulaniemi, J. (2011). *Palvelumuotoilu*. Karisto. Talentum Media Oy.

Tuulenmäki, A. and Välikangas, L. (2011). *The Art of Rapid, Hands-on Execution Innovation*. Strategy & Leadership 39.2, pp. 28–35.

Tversky, A. and Kahneman, D. (1974). *Judgment Under Uncertainty: Heuristics and Biases*. Science, New Series, Vol. 185, No. 4157.

Tversky, A. and Kahneman, D. (1983). *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*. Psychological Review 90, 293–315.

USPSTF (1989). *Guide to Clinical Preventive Services: Report of the U.S. Preventive Services Task Force*, p. 24–. August 1989. U.S. Preventive Services Task Force. DIANE Publishing.

Vaajakallio, K. (2012). *Design Games as a Tool, A Mindset and A Structure*, p. 14. Aalto University publication series. Doctoral Dissertations 87/2012. Aalto University School of Arts, Design and Architecture. Helsinki: Aalto ARTS Books.

Vaajakallio, K. and Mattelmäki, T. (2013). *Co-design with the Public Sector*, p. 103-115. In: K. Vaajakallio and J. Honkonen (eds.). Designing for Wellbeing. Helsinki: Aalto ARTS Books.

Vaajakallio, K. Lee, J-J., Kronqvist, J. and Mattelmäki, T. (2013). *Service Co-Design with the Public Sector - Challenges and Opportunities in a Healthcare Context*. Include Asia 2013 Proceedings. London: Helen Hamlyn Center RCA.

Voutier, C. (2014). *What Are Hierarchies of Evidence and Why Are They Used?* HLinc Presentation. Health Sciences Library. May 2014. Retrieved 20 Feb. 2017 from <http://www.slideshare.net/Catherine Voutier/hl-presentation-levels-of-evidence>

Vuontisjärvi, H-R. (2005). *Service Design In The Public Sector: Role, Challenge and Good Practice*, p. 47. University of Lapland. Retrieved 20 Feb. 2017 from <http://lauda.ulapland.fi/bitstream/handle/10024/ 62022/pro+Gradu_Vuontisj%C3%A4rvi_2015.pdf?sequence=2>

Whaites, A. (2008). *States in Development: Understanding State-Building*, p. 9. A DFID Working Paper. Department for International Development.

Wheatley, M. J. and Kellner-Rogers, M. (2012). *A Simpler Way*. Berrett-Koehler Publishers.

Whicher, A. (2015). *The Value of Design to the Public Sector. Design For Europe*. Retrieved 20 Feb. 2017 from <http://designforeurope.eu/news-opinion/value-design-public-sector>

Wikipedia (2007). *List of cognitive biases*. Wikipedia. The Free Encyclopedia. Retrieved 20 Feb. 2017 from <https://en.wikipedia.org/wiki/List_of_cognitive_biases>

Williams-Grut, O. (2014). *London launch for 'D5' alliance of digital nations*. The Independent. Retrieved 20 Feb. 2017 from <http://www.independent.co.uk/news/business/news/london-launch-for-d5- alliance-of-digital-nations-9909374.html>

Wright, O. (2014). *'The Nudge team started out as a sort of Mission Impossible': How the Government's successful Behaviour Insights Team has had a profound effect on Whitehall.* The Independent. Retrieved 20 Feb. 2017 from <http://www.independent.co.uk/news/people/profiles/the-nudge-team-started-out- as-a-sort-of-mission-impossible-how-the-government-s-successful-behaviour-9117793.html>

Yee, J. and Choukeir, J. (2016). *From Capability Training to Capacity Building*, p. 68-73. In: B. Mager (ed.). Service Design Impact Report: Public Sector. Köln: Service Design Network. Retrieved 20 Feb. 2017 from <https://www.service-design-network.org/books- and-reports/impact-report-public-sector>

Yeung, K. (2012). *Nudge as Fudge*, p. 4. Modern Law Review. 75(1):122–148.

ANNEX 1

Sample Interview Request E-mail

Dear (Recipient Name),

My name is Ilari Laitinen and I am a Master's student from Aalto University in Finland, with International Design Business Management (IDBM) as my major. I am currently working on my Master's Thesis on the subject of 'Evidence-based Practices in Public Service Design'. My research focuses currently on evidence-based practices to evaluate the effectiveness of public services. The central research question in my thesis is based on how service designers can use evidence to support the design process. The aim of the study is to strengthen the knowledge base of the service design to foster the implementation of these practices so they can be used in a practical design work to ensure meaningful public services and contributing to the issue of the service design's legitimacy and codes of practice.

My question is: will you be willing to be interviewed for my thesis?

(Consultancy Name) is a great example of an consultancy that is promoting the use of design methods to support the development of programmes and policies within the public sector. I strongly believe that you have valuable information, experience and vision regarding this Master's Thesis. I would like to hear more about your thoughts, practices and experience in this field.

My intention is to interview service designers as well as different experts from the fields of public policy, academia and behavioural economics in Finland, Sweden, United Kingdom and the United States all through March 2016. However, I'm aware that you might be quite busy, so I assure you I will be brief and concise – taking up no more than 1-2 hours of your time. The interview could be either through a phone call, or by Skype.

Thank you very much for considering this request. I look forward to receiving a response from you.

ANNEX 1

Sample Interview Questions (*Group A*)

Q1: How would you define a successful service design project in the public sector?

Probe: Would you give some examples of the projects that you have done for the public sector?

Q2: How would you describe the difference in approach between your public sector and private sector projects?

Sub-question: What are the opportunities involved in working with the public sector in (your country)?
Sub-question: What are the challenges involved in working with the public sector in (your country)?

Probe: What is the procedure for starting a public sector service project in (your country)? Is there typically a call for bids, or call for tenders, where the buying organisation has specified the requirements for the service design project?

Probe: How is the relationship between policy-makers and service designers in (your country)?

Probe: Do they typically understand the scope of service design and its value for public service delivery?

Probe: What is the role of a service designer in the public sector projects?

Q3: How would you characterise or categorise the different public service design projects?

Probe: How would you characterise a typical high level (or strategic) project for the public sector?

Probe: How do these two ends differ from each other?

Q4: What methods do you typically use in the public service design projects to collect customer insights or research data?

Probe: What kind of data are you typically looking for that supports your decision-making?

Probe: Are there any challenges in gathering this data?

Probe: Is there a need for more rigid research methods?

Q5a: Do you conduct experiments or interventions to test your ideas or hypotheses?

Probe: What are the important things to take into consideration when designing these experiments?

Probe: How long do they usually last? What kind of data do you expect to produce?

Probe: What are the situations that will likely cause the experiment to fail?

Q5b: Can you give some examples of the experiments that you have done with the public sector?

Sub-question: How do you evaluate the experiments?

Probe: How do you measure the experiment outcomes in terms of changes in desired outcome indicators before and after the experiments?

Probe: Do you set desired outcome indicators before experiments? How these are defined?

Probe: Do you obtain data using control groups or other more rigid research methods?

Q6: How is your work evaluated during and after the public service design projects?

Probe: What are the problem owners typically looking for in these reviews?

Probe: At what stage of the process do they typically review your work?

Probe: How do you justify your decisions in the design process?


A!
Aalto University